

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 29, 2005, 23:18:18 ; Search time 2993.55 Seconds  
(without alignments)  
1990.942 Million cell updates/sec

Title: US-10-657-740-1\_COPY\_51\_173

Perfect score: 639

Sequence: 1 SLFRTVLDGISEVSRDRK.....HAERAIPVSRBKPTSAFSS 123

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US10657740/runat 27052005 165253 3315/app query.fasta\_1.590  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=150  
-DOCALIGN=200 -NORM=scorepct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10657740 @CGN 1 1 4362 @runat 27052005 165253 3315 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sv.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	639	100.0	519	9	CR407691 Homo sapi
2	639	100.0	614	9	BC069528 Homo sapi
3	639	100.0	741	9	U66584 Human alpha
4	639	100.0	1112	9	U05569 Human alpha

5	639	100.0	1114	6	CQ731849
6	602	94.2	543	4	OCRYAA
7	599	93.7	1025	10	MUSCRYAB
8	599	93.7	1056	10	RNU47922
9	598	93.6	793	4	BOVCYA
10	591	92.5	885	10	RNACHY
11	577.5	90.4	796	10	MMU310308
12	577.5	90.4	1271	10	RNU47921
13	550	86.1	447	4	AJ617725
14	548	85.8	477	5	AJ617726
15	544	85.1	522	5	RCACRYST
16	528.5	82.7	558	4	AJ617724
17	519	81.2	765	5	RVCRYA
18	507	79.3	1361	5	D86299
19	505	79.0	447	5	AJ617727
20	502	78.6	448	5	TSU31938
21	492	77.0	705	5	D88185
22	487	76.2	448	5	EUDLACRY5
23	486	76.1	448	5	APRAAC
24	476	74.5	448	5	CLRAAC
25	474	74.2	730	5	AY035778
26	474	74.2	1430	5	BC083177
27	466	72.9	676	5	AY007972
28	408	63.8	436	5	OLJ000940
29	404	63.2	249	10	RATACRYA
30	401	62.8	249	10	RATACRYB
31	367	57.4	419	9	AF026952
32	367	57.4	148179	9	AP001631
33	367	57.4	163962	9	BS000233
34	367	57.4	340000	9	AP001748
35	364	57.0	507	4	AJ617732
36	364	57.0	632	4	AF029793
37	361	56.5	671	10	S74229
38	361	56.5	687	10	RRHARTABC
39	361	56.5	689	10	RRELENSABC
40	361	56.5	704	10	S77142
41	361	56.5	706	10	S77138
42	361	56.5	1247	6	AX401743
43	361	56.5	1247	10	RATCRYAB
44	358	56.0	548	4	OCRYAB
45	357	55.9	511	10	SEH272441
46	357	55.9	528	10	SJU293658
47	356	55.7	537	5	RCACRYST
48	356	55.7	537	10	HAMSCRYB
49	354	55.4	528	9	BT006770
50	354	55.4	528	12	BT007909
51	354	55.4	537	6	AX937702
52	354	55.4	537	6	AX937703
53	354	55.4	691	6	CQ812349
54	354	55.4	691	6	AX330255
55	354	55.4	691	6	AX330322
56	354	55.4	691	6	BD079402
57	354	55.4	691	9	S45630
58	354	55.4	744	9	BC007008
59	354	55.4	856	6	AX888028
60	354	55.4	856	6	BD027638
61	354	55.4	893	6	AX899079
62	354	55.4	893	6	BD034612
63	354	55.4	911	6	AX899075
64	354	55.4	911	6	BD034608
65	354	55.4	913	6	AX899074
66	354	55.4	913	6	BD034607
67	354	55.4	927	6	AX899073
68	354	55.4	927	6	BD034606
69	354	55.4	942	6	AX888027
70	354	55.4	942	6	BD027637
71	354	55.4	1036	6	BD135185
72	354	55.4	1036	6	AX017495
73	353	55.2	733	9	AB125159
74	352	55.1	666	10	MUSCRYABA
75	352	55.1	848	10	BC010768
76	352	55.1	966	5	BC082122
77	350	54.8	661	9	AF007162

c

c

c

78	343	53.7	689	5	DUKABC	L08078	Anas platyr
79	343	53.7	906	10	MAACRY2	X02951	Hamster alp
80	342	53.5	846	5	GGU26661	U26661	Gallus gall
81	342	53.5	1042	5	BX930014	BX930014	Gallus ga
82	342	53.5	1251	5	S53164	S53164	alpha B-cry
83	340	53.2	167475	10	AC090881	AC090881	Mus Muscu
84	340	53.2	188810	2	AC121138	AC121138	Mus Muscu
85	340	53.2	222895	2	AC084065	AC084065	Mus Muscu
86	339	53.1	466	4	AJ617731	AJ617731	Macropus
87	339	53.1	900	5	AF159089	AF159089	Danio rer
88	332	52.0	466	4	AJ617730	AJ617730	Diolephis
89	332	52.0	502	4	AJ617729	AJ617729	Tachyglis
90	330	51.6	577	5	AY007973	AY007973	Clarias b
91	330	51.6	695	6	AX888029	AX888029	Sequence
92	330	51.6	695	6	BD027639	BD027639	Sequence
93	329	51.5	501	4	AJ617728	AJ617728	Ornithorh
94	321	50.2	1316	10	SEHCYAA2	MI7249	Mole rat al
95	320	50.1	1955	5	BC076518	BC076518	Danio rer
96	317.5	49.7	765	5	RTCRYA	X00716	Frog mRNA f
97	310.5	48.6	888	5	BC075197	BC075197	Xenopus l
98	310	48.5	6311	5	CHKCRYAA	MI7627	Chicken alp
99	309	48.4	826	6	AR415535	AR415535	Sequence
100	309	48.4	826	6	AX972369	AX972369	Sequence
101	309	48.4	826	6	BD111088	BD111088	EST and e
102	293	45.9	168	6	HUMAAAC	L25781	Homo sapien
103	282.5	44.2	373	6	CQ681472	CQ681472	Sequence
104	275.5	43.1	1065	5	AY184812	AY184812	Clarias b
105	273.5	42.8	383	11	BV209856	BV209856	CRYAA.464
106	271	42.4	830	3	AF237691	AF237691	Ciona Int
107	268	41.9	787	10	CLSHSP	X51747	Cricetulus
108	267	41.8	763	5	GGIAP	X59541	Chicken mRN
109	266.5	41.7	165693	5	BX248514	BX248514	Zebrafish
110	266.5	41.7	166831	2	BX950201	BX950201	Danio ter
111	263	41.2	349	6	AX786923	AX786923	Sequence
112	263	41.2	882	10	MUSHSP25PS	Li1610	Mus musculu
113	262	41.0	1085	5	BC078509	BC078509	Xenopus l
114	261.5	40.9	615	9	CR407614	CR407614	Homo sapi
115	261.5	40.9	618	9	CR536489	CR536489	Homo sapi
116	261.5	40.9	640	9	HSW800252	AL050380	Homo sapi
117	261.5	40.9	724	9	BC014920	BC014920	Homo sapi
118	261.5	40.9	764	6	CQ799993	CQ799993	Sequence
119	261.5	40.9	764	9	AB020027	AB020027	Homo sapi
120	261.5	40.9	781	9	BC012768	BC012768	Homo sapi
121	261.5	40.9	789	6	AR380755	AR380755	Sequence
122	261.5	40.9	789	9	HSRSP27L	X54079	Human mRNA
123	261.5	40.9	794	9	BC073768	BC073768	Homo sapi
124	261.5	40.9	847	6	BD186273	BD186273	STAT6 act
125	261.5	40.9	865	6	CQ730135	CQ730135	Sequence
126	261.5	40.9	865	9	HSU90906	U90906	Human clone
127	261.5	40.9	867	9	BC000510	BC000510	Homo sapi
128	261.5	40.9	1231	6	AX411221	AX411221	Sequence
129	261.5	40.9	1231	9	HSRSP28	Z23090	H. sapiens m
130	261.5	40.9	1380	6	AR217508	AR217508	Sequence
131	261.5	40.9	1380	6	AX401305	AX401305	Sequence
132	261	40.8	604	6	AX401305	U03562	Mus musculu
133	260	40.7	535	10	MMU03562	U03561	Mus musculu
134	260	40.7	601	6	MMU03561	A76413	Sequence 1
135	260	40.7	634	10	MMU03560	U03560	Mus musculu
136	260	40.7	787	6	CQ858777	CQ858777	Sequence
137	260	40.7	787	6	AX401752	AX401752	Sequence
138	260	40.7	787	6	RATHSP27A	M86389	Rattus norv
139	260	40.7	862	10	BC018257	BC018257	Mus muscu
140	260	40.7	862	10	AR175769	AR175769	Sequence
141	260	40.7	1379	6	AR232004	AR232004	Sequence
142	260	40.7	1379	6	AX136008	AX136008	Sequence
143	260	40.7	1379	6	AX136777	AX136777	Sequence
144	260	40.7	1379	6	AX137828	BD000152	Process f
145	260	40.7	1379	6	BD000152	BD000152	Process f
146	260	40.7	1379	6	BD010884	BD010884	Process f
147	260	40.7	1379	6	BC078511	BC078511	Xenopus l
148	259.5	40.6	832	5	BC078511	BC078511	Xenopus l
149	259.5	40.6	4321	5	APACCEPIG	Y11300	Astyanax fa
150	259	40.5	228912	2	AC114562	AC114562	Mus muscu

## ALIGNMENTS

RESULT 1	CR407691	519 bp	mRNA	linear	PRI 10-MAY-2004
LOCUS	CR407691				
DEFINITION	Homo sapiens full open reading frame cDNA clone RZPD0834E043D for gene CRYAA, crystallin, alpha A complete cds, without stopcodon.				
ACCESSION	CR407691				
VERSION	CR407691.1	GI:47115318			
KEYWORDS	Full ORF shuttle clone, Gateway(TM), complete cds.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Ebert,L., Schick,M., Neubert,P., Schatten,R., Henze,S. and Korn,B.				
TITLE	Cloning of human full open reading frames in Gateway(TM) system entry vector (pDONR201)				
JOURNAL	Unpublished				
AUTHORS	2 (bases 1 to 519)				
TITLE	Ebert,L., Schick,M., Neubert,P., Schatten,R., Henze,S. and Korn,B.				
JOURNAL	Direct Submission				
COMMENT	Submitted (07-MAY-2004) RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany RZPD; RZPD0834E043D, ORFNO 642 www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPD0834E043D RZPDLIB; Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No. 834 www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834 www.rzpd.de/products/orfclones/ Contact: Ina Rolfs RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Tel: +49 30 32639 101 Fax: +49 30 32639 111 www.rzpd.de This clone is available from RZPD; contact RZPD (customer.service@rzpd.de) for further information. This CDS clone is a part of a collection of human full length expression clones generated by RZPD. This CDS has been inserted into pDONR201 via a BP Clonase(TM) reaction. Additional sequence has been added in front of the start codon (ATG): att...AAAAA GCT GGC ACC CCT GGT CCA GGT (ATG) After the last codon additional sequence has been added: CCA GGC CCA GGC GGC G in front of the 3' att site (AC CCA GCT TTC TT). Compared to the reference sequence U66584 we did not find any amino acid exchanges. Clone distribution: http://www.rzpd.de/products/orfclones/. FEATURES source Location/Qualifiers 1..519 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="RZPD0834E043D" /clone_lib="Human Full ORF Clones Gateway(TM) - RZPD" /lab_host="DH10B" /note="Vector: pDONR201, Site_1: attP1; Site_2: attP2" 1..519 /gene="CRYAA" 1..519 /gene="CRYAA" /codon_start=1 /protein_id="CAG28619.1" /db_xref="GI:47115318" SPYRQSLFRTVLDGSEVSDRDKFVFLDVKHFSPEDLTVKQDDFVEIHGKNE RODDHGYSREHFHRYRLPSNVNDSALSCSLSDGMLTFCGPKIQTGLDTHAERAIP VSREKPTSPSS"				
gene	CRYAA				
CDS	1..519				
ORIGIN	CRYAA				

Alignment Scores:	3.77e-63	Length:	519
Pred. No.:	639.00	Matches:	123
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-10-657-740-1_COPY_51_173 (1-123) x CR407691 (1-519)			
Qy	1	1	1
Db	151	151	151
Qy	21	21	21
Db	211	211	211
Qy	41	41	41
Db	271	271	271
Qy	61	61	61
Db	331	331	331
Qy	81	81	81
Db	391	391	391
Qy	101	101	101
Db	451	451	451
Qy	121	121	121
Db	511	511	511
RESULT 2	BC069528	614 bp	linear
LOCUS	BC069528	IMAGE:7262133	alpha A, mRNA (cdna clone MGC:96924
DEFINITION	BC069528	complete cds.	
ACCESSION	BC069528	GI:46854598	
VERSION	BC069528.1		
KEYWORDS	MGC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1	(Bases 1 to 614)	
AUTHORS	Strausberg, R.L., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheerz, T.E., Brownstein, M.J., Usdin, T.B., Toehiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, U., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		

12477932	PUBMED	2	(Bases 1 to 614)
Strausberg, R.	REFERENCE		
Submitted (29-APR-2004)	AUTHORS		
Gene Collection (MGC), Cancer Genomics Office, National Cancer	TITLE		
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,	JOURNAL		
USA			
NIH-MGC Project URL: http://mgc.nci.nih.gov			
Contact: MGC help desk			
Email: cgapbs-remail.nih.gov			
Tissue Procurement: Baylor Human Genome Sequencing Center			
cDNA Library Preparation: Baylor Human Genome Sequencing Center			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
DNA Sequencing by: Baylor College of Medicine Human Genome			
Sequencing Center			
Center code: BCM-HGSC			
Web site: http://www.hgsc.bcm.tmc.edu/cdna/			
Contact: amgobcm.tmc.edu			
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleeged, H.,			
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,			
A.N., Gibbs, R.A.			
Clone distribution: MGC clone distribution information can be found			
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov			
Series: IRBR Plate: 1 Row: d Column: 7.			
Location/Qualifiers			
1. 614			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="MGC:96924 IMAGE:7262133"			
/tissue type="PCR rescued clones"			
/clone lib="NIH MGC 244"			
/note="vector: pPCR-Script Amp SK(+)"			
1. 614			
/gene="CRYAA"			
/note="synonyms: CRYA1, HSPB4"			
/db_xref="LocusID:1409"			
/db_xref="MIM:123580"			
51-572			
/gene="CRYAA"			
/codon_start=1			
/product="crystallin, alpha A"			
/protein id="AAH69528.1"			
/db_xref="GI:46854599"			
/db_xref="LocusID:1409"			
/db_xref="MIM:123580"			
/translation="MDVTIQHPFKRTLPFPYPSRLPDPQFGEGLFYEDLLPFLSSTI			
SPYRQSLFRFLVDSGISSEVRSDRDKFVFLDVKHFSPEDLVKQDDFVEIHGHNE			
RQDDHGVISREFHRRYRLPSNVDSALSCSLSDGMLTFCGPKIQTLGLDTHAERAIP			
VSREKPTSPSS"			
ORIGIN			
Alignment Scores:			
Pred. No.:	4.59e-63	Length:	614
Score:	639.00	Matches:	123
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-10-657-740-1_COPY_51_173 (1-123) x BC069528 (1-614)			
Qy	1	1	1
Db	201	201	201
Qy	21	21	21
Db	261	261	261
Qy	41	41	41

```
Db 321 GACGACTTTGTGGAGATCCAGGAAGCACACGAGAGCCAGGACGACCGGCTACATT 380
Qy 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 381 TCCCGTGAAGTTCCACCGCGCTACCGCTCCGCTCCAAACGTCGACGACGCTCTCTCT 440
Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 441 TCTCTCTCTGTGCGCATGGCATGCTGACCTTCTGTGCGCCCAAGATCCAGCTGGCGCTG 500
Qy 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 501 GATGCCACCCACGCGAGCGAGCCATCCCGTGTGCGGGAGGAGAGCCACCTCGGCT 560
Qy 121 ProSerSer 123
Db 561 CCCTCGTCC 569

RESULT 3
HSU6584
LOCUS Human alphaA-crystallin (CRYAA) mRNA, complete cds. PRI 14-DEC-1996
DEFINITION
ACCESSION U66584
VERSION U66584.1 GI:1732062
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Petrash,J.M., Mathur,S., Manoharan,M. and Andley,U.P.
TITLE Cloning and expression of human lens crystallins
JOURNAL Invest. Ophthalmol. Vis. Sci. 36, S882-S882 (1995)
REFERENCE
AUTHORS Andley,U.P., Mathur,S., Griest,T.A. and Petrash,J.M.
TITLE Cloning, expression, and chaperone-like activity of human
alphaA-crystallin
JOURNAL J. Biol. Chem. 271 (50), 31973-31980 (1996)
MEDLINE 97112991
PUBMED 8943244
REFERENCE
AUTHORS Petrash,J.M., Mathur,S., Wang,J.C., Griest,T.A. and Andley,U.P.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1996) Ophthalmology and Visual Sciences,
Washington University School of Medicine, 660 S. Euclid Ave., St.
Louis, MO 63110, USA

FEATURES
source
location/Qualifiers
1..741
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="eye lens"
1..741
/gene="CRYAA"
1..522
/gene="CRYAA"
/function="structural protein"
/note="structurally related to sHSPs; chaperone-like
activity"
/codon_start=1
/product="alphaA-crystallin"
/protein_id="AAC5090.1"
/db_xref="GI:1732063"
/translation="MDVTIQHFWFKTLGPFVPSRLFDQFFGEGLEFYEDLLPFLSSTI
SPYKSLRFTVLDSISVSRDRKPFVLDVKHFSPEDLTVKQDDFVEIHGKNE
RQDHGYISREPHRYRLFSNVDQSALSCSLADGMLTCGPKIQTLGDATHAERAIP
VSREKPTSPSS"
```

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x HSU6584 (1-741)

Qy 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 151 TCCCTCTTCGCGACCGTGTGGACTCGGCGATCTCTGAGGTTGATCCGACCGGAGCAAG 210
Qy 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40
Db 211 TTCGTCTATCTTCTCGATGTGAAGCACTTCTCCCGGAGGACCTCACCGTGAGGTGCAG 270
Qy 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60
Db 271 GACGACTTTGTGGAGATCCACGGAAGACACACGAGCGCCAGGACGACCGGTACATT 330
Qy 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 331 TCCCGTGAAGTTCCACCGCGCTACCGCTCCGCTCCAAACGTCGACGACGCTCTCTCT 390
Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 391 TGCTCCCTGTGTCGCGATGGCATGCTGACCTTCTGTGCGCCCAAGATCCAGACTGGCGTG 450
Qy 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 451 GATGCCACCCACGCGAGCGAGCCATCCCGTGTGCGGGAGGAGAGCCACCTCGGCT 510
Qy 121 ProSerSer 123
Db 511 CCCTCGTCC 519

RESULT 4
HSU05569
LOCUS Human alphaA-crystallin (CRYA1) mRNA, complete cds. PRI 25-APR-1996
DEFINITION
ACCESSION U05569
VERSION U05569.1 GI:452477
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Jaworski,C.J.
TITLE The human alphaA-crystallin gene
JOURNAL Thesis (1992) LMDB, NEI, Molecular Structure and Function
REFERENCE 2 (bases 1 to 1112)
AUTHORS Jaworski,C.J.
TITLE A reassessment of mammalian alpha A-crystallin sequences using DNA
sequencing: implications for anthropoid affinities of tarsier
J. Mol. Evol. 41 (6), 901-908 (1995)
MEDLINE 96139023
PUBMED 8587135
REFERENCE 3 (bases 1 to 1112)
AUTHORS Wistow,G.J.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-1994) Graeme J. Wistow, Molecular Structure and
Function, LMDB, NEI, NIH, Bethesda, MD 20892, USA

FEATURES
source
location/Qualifiers
1..1112
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/tissue_type="lens"
1..1112
/gene="CRYA1"
68..589
/gene="CRYA1"
```

```
Alignment Scores: 5.72e-63 Length: 741
Pred. No.: 639.00 Matches: 123
Score:
```



/function="lens structural protein"  
 /note="This sequence is derived from genomic and cDNA  
 sequence"  
 /citation={1}  
 /citation={2}  
 /codon\_start=1  
 /evidence=experimental  
 /product="alphaA-crystallin"  
 /protein\_id="AA097523.1"  
 /db\_xref="GI:452478"  
 /translation="MDVTIQHPFKRTGLGPPSRFLDQFEGEGLEFYDLPLFLSSTI  
 SPYKQSLPRTVDSGISVRSRDKFVIFLDVKHSPEDLTIVQDDFVHKGHNE  
 RQDDGXTLSRFHRYRLPSNVDSALSLSADGMLTFCGPKIQTLGLDTHAERAIP  
 VSREKPTSAFSS"

## ORIGIN

Alignment Scores:  
 Pred. No.: 9,19e-63 Length: 1112  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x HSU05569 (1-1112)

Qy 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 218 TCCCTCTTCCGACCGTCTGACCTCCGGCATCTCTGAGGTTCCGATCCGACCGGACAA 277  
 Qy 21 PheValIlePheLeuAspValIlyshisPheSerProGluAspLeuThrValIlyValGln 40  
 Db 278 TTGCTCATCTTCTCTCGATGTGAAGCACCTTCTCCCGGAGGACCTCACCGTGAAGGTGCAG 337  
 Qy 41 AspAspPheValGluIleHisGlyIlyshisGluArgGlnAspAspHisGlyTyrlle 60  
 Db 338 GACGACTTTGTGGAGATCCAGAAAGCACACACGACGCGGACGACGACGCTACATT 397  
 Qy 61 SerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 398 TCCCGTGAAGTTCACCGCGCTACCGCTCCGCTCCACGTTGACCATCGGCGCTCTCT 457  
 Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 Db 458 TGCTCCCTGTCTGCGGATGCGATGCTGACCTTCTGTGCGCCCAAGATCCAGACTGCGCTG 517  
 Qy 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 518 GATGCCACCCACGCGGAGGACCATCCCGTCCGCTCCGCGGAGGAGAGGCCACCTCGGCT 577  
 Qy 121 ProSerSer 123  
 Db 578 CCCTCGTCC 586

RESULT 5  
 LOCUS COQ731849 1114 bp DNA linear PAT 03-FEB-2004  
 DEFINITION Sequence 17783 from Patent WO2068579.  
 ACCESSION COQ731849  
 VERSION COQ731849.1 GI:42309516  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
 TITLE Kits, such as nucleic acid arrays, comprising a majority of  
 humanexons or transcripts, for detecting expression and other uses  
 thereof  
 JOURNAL Patent: WO 02068579-A 17783 06-SEP-2002;  
 PE Corporation (NY) (US)  
 FEATURES Location/Qualifiers

## source

1. .1114  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:  
 Pred. No.: 9,21e-63 Length: 1114  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x CQ731849 (1-1114)

Qy 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 220 TCCCTCTTCCGACCGTCTGACCTCCGGCATCTCTGAGGTTCCGATCCGACCGGACAA 279  
 Qy 21 PheValIlePheLeuAspValIlyshisPheSerProGluAspLeuThrValIlyValGln 40  
 Db 280 TTGCTCATCTTCTCTCGATGTGAAGCACCTTCTCCCGGAGGACCTCACCGTGAAGGTGCAG 339  
 Qy 41 AspAspPheValGluIleHisGlyIlyshisGluArgGlnAspAspHisGlyTyrlle 60  
 Db 340 GACGACTTTGTGGAGATCCAGAAAGCACACACGACGCGGACGACGCTACATT 399  
 Qy 61 SerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 400 TCCCGTGAAGTTCACCGCGCTACCGCTCCGCTCCACGTTGACCATCGGCGCTCTCT 459  
 Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 Db 460 TGCTCCCTGTCTGCGGATGCGATGCTGACCTTCTGTGCGCCCAAGATCCAGACTGCGCTG 519  
 Qy 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 520 GATGCCACCCACGCGGAGGACCATCCCGTCCGCTCCGCGGAGGAGAGGCCACCTCGGCT 579  
 Qy 121 ProSerSer 123  
 Db 580 CCCTCGTCC 588

## RESULT 6

LOCUS OCCRYAA 543 bp mRNA linear MAM 23-OCT-1996  
 DEFINITION O.cuniculus mRNA for alpha-A-crystallin.  
 ACCESSION X95382  
 VERSION X95382.1 GI:1177576  
 KEYWORDS alpha-A-crystallin; cryaA gene.  
 SOURCE Oryctolagus cuniculus (rabbit)  
 ORGANISM Oryctolagus cuniculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 REFERENCE  
 AUTHORS Krausz,E., Augusteyn,R.C., Quinlan,R.A., Reddan,J.R., Russell,P.,  
 Sax,C.M. and Graw,J.  
 TITLE Expression of Crystallins, Pax6, Filensin, CP49, MIP, and MP20 in  
 lens-derived cell lines  
 JOURNAL Invest. Ophthalmol. Vis. Sci. 37 (10), 2120-2128 (1996)  
 MEDLINE 96409169  
 PUBMED 8814151  
 REFERENCE 2 (bases 1 to 543)  
 AUTHORS Graw,J.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-JAN-1996) J. Graw, Institute of Mammalian Genetics,  
 GSF-Research Center Neuherberg, Ingolstaedter Landstrasse 1,  
 Oberschleissheim, D-85764, FRG  
 FEATURES Location/Qualifiers  
 source 1. .543  
 /organism="Oryctolagus cuniculus"  
 /mol\_type="mRNA"

```

/sub_species="New Zealand white"
/db_xref="taxon:9986"
/tissue_type="lens"
1..543
/gene="cryaa"
22..543
/gene="cryaa"
/codon_start=1
/product="alpha-A-crystallin"
/protein_id="CAA64668.1"
/db_xref="GI:1177577"
/db_xref="GOA:P02493"
/db_xref="UniProt/Swiss-Prot:P02493"
/translation="MDVTIOHPFKRTIGPPYPSRLFDQFFGEGFLFEYDILLPLFSLSTI
SPYRQSLFRTVLDGSISEVRSDRKDFVFLDVKHFSPEDLTVKVOEDFVEIHGKHNE
RODDHGYSREHRRYRLPSNVDSALSCSLSDGMLTFSGPKVQSGLDAGHSERAI
VSREKPSVPSS"
ORIGIN
Alignment Scores:
Pred. No.: 6.27e-59 Length: 543
Score: 602.00 Matches: 115
Percent Similarity: 97.56% Conservative: 5
Best Local Similarity: 93.50% Mismatches: 3
Query Match: 94.21% Indels: 0
DB: 4 Gaps: 0
US-10-657-740-1_COPY_51_173 (1-123) x OCCRYAA (1-543)
QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 172 TCCCTCTTCCGACCGCTGCTGACTCCGGCATCTCTGAGTCCGCTCCGACGGGACAA 231
QY 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40
Db 232 TTGCGTCATCTTCTGAGCGTGAAGACACTTCTCCGCGGAGGACCTCACCGTGAAGTGCAG 291
QY 41 AspAspPheValGluIleHisGlyIleHisGlyIleHisGlyIleHisGlyIleHisGlyIle 60
Db 292 GAGGACTCTGAGATCCACGGCAACACACAGAGACAGATGACCAAGCTACATT 351
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 352 TCCCGCGAGTTCCACGGCGCTACCGCTGCTCCCTCCAACTGCAACAGTCCGCGCTGCTCC 411
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 412 TGCTCGCTCTCCGCGACGGCATGCTACCTTCTCCGGCCCCCAAGTGAGTCCGCGCTG 471
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 472 GACGCTGGCCACAGCGAGCGACCCATCCCGTGTCCCGGAGGAGAGCCAGCTCGGTG 531
QY 121 ProSerSer 123
Db 532 CCCTCGTCC 540
RESULT 7
MUSCRYAAB 1025 bp mRNA linear ROD 27-APR-1993
LOCUS Mouse lens alpha-A-crystallin mRNA, 3' end.
DEFINITION J00376
ACCESSION J00376.1 GI:192760
VERSION alpha-A-crystallin; crystallin.
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1025)
AUTHORS King, C.R., Shinohara, T. and Piatigorsky, J.
TITLE alpha-A-crystallin messenger RNA of the mouse lens: more noncoding
than coding sequences
JOURNAL Science 215 (4535), 985-987 (1982)

```

```

MEDLINE 83119896
PUBMED 7156978
COMMENT
Original source text: Mouse lens, cDNA to mRNA.
The crystallins contribute 90% of the soluble lens protein and are
highly conserved during evolution. There are four immunologically
separate classes of crystallins called alpha-, beta-, gamma-, and
delta-crystallin. The mRNA of the alpha-A-crystallin from mouse,
like that from cows and rats, is nearly three times the size
required to code for the polypeptide. Nucleotides 1-490 of the
mouse alpha-A-crystallin mRNA predict a protein sequence identical
to amino acids 10-173 of rat alpha-A-crystallin.
FEATURES
Location/Qualifiers
source
1..1025
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
<1..1025
/product="CRY-alpha-A mRNA"
<1..492
/notes="alpha-A-crystallin"
/codon_start=1
/protein_id="AAA37471.1"
/db_xref="GI:387134"
/translation="KRALGPFYPSRLFDQFFGEGFLFEYDILLPLFSLSTI
TVLDGSISEVRSDRKDFVFLDVKHFSPEDLTVKVOEDFVEIHGKHNERODDHGYSR
EFHRRYRLPSNVDSALSCSLSDGMLTFSGPKVQSGLDAGHSERAI
PSS"
ORIGIN 70 bp upstream of Real site.
Alignment Scores:
Pred. No.: 2.89e-58 Length: 1025
Score: 599.00 Matches: 115
Percent Similarity: 97.56% Conservative: 5
Best Local Similarity: 93.50% Mismatches: 3
Query Match: 93.74% Indels: 0
DB: 10 Gaps: 0
US-10-657-740-1_COPY_51_173 (1-123) x MUSCRYAAB (1-1025)
QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 121 TCCCTCTTCCGACCTGCTGACTCCGGCATCTCTGAGTCCGCTCCGACGGGACAA 180
QY 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40
Db 181 TTGTCATCTTCTGAGCGTGAAGACACTTCTCTCTGAGGACCTCACCGTGAAGTACTG 240
QY 41 AspAspPheValGluIleHisGlyIleHisGlyIleHisGlyIleHisGlyIleHisGlyIle 60
Db 241 GAAGATTGTTGGAGATTACGGCAACACACAGAGGAGGAGGATGACCTGCGCTACATA 300
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 301 TCCCGTGAATTTCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 361 TGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 421 GATGCTGGCCACAGCGAGGAGGCGCATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 121 ProSerSer 123
Db 481 CCCTCGTCC 489
RESULT 8
LOCUS RNU47922
DEFINITION Rattus norvegicus alpha A-crystallin mRNA, complete cds.
ACCESSION U47922
VERSION U47922.1 GI:1245161

```

**KEYWORDS**  
**SOURCE** Rattus norvegicus (Norway rat)  
**ORGANISM** Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
**REFERENCE** 1 (bases 1 to 1056)  
**AUTHORS** Bhat, S.P., Nandy, P., Srinivasan, A., Cheng, D. and Sitay, A.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (31-JAN-1996) Suraj P. Bhat, JSEI, UCLA School Medicine,  
 100-Stein Plaza, Rm. BH623, Los Angeles, CA 90095-7008, USA  
**FEATURES**  
 Location/Qualifiers  
 source  
 1..1056  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 13..534  
 /note="similar to small heat shock proteins"  
 /codon\_start=1  
 /product="alpha A-crystallin"  
 /protein\_id="AA93367.1"  
 /db\_xref="GI:1245162"  
 /translation="MDVTHQHPWFKRALGPVPSRLFOQFGEGLFEYDLPLFLSSTI  
 SPYRQSLFRVLDSGISVSRDRDKFVFLDVKHFSPELITVKVLEDFVEIHGKNE  
 RQDDHGYSIRSEFHRRYRLPSNVDSQALSCLSDGMLTFSKPKVQSGLDAGHSRAIP  
 VSREKFPSSAPSS"  
 535..1050  
**3'UTR**  
**ORIGIN**

Alignment Scores:  
 Pred. No.: 2,996-58 Length: 1056  
 Score: 599.00 Matches: 115  
 Percent Similarity: 97.56% Conservative: 5  
 Best Local Similarity: 93.50% Mismatches: 3  
 Query Match: 93.74% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x RNU47922 (1-1056)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 DB 163 TCTCTTCCGACAGTGTGGACTCCGGCATCTCTGAGTCCGATCTGACCGGACAG 222  
 QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
 DB 223 TTTGTCTCTTCTGGATGTGAAGCACCTCTCTCTGAGGACCTCACCGTGAAGTACTG 282  
 QY 41 AspAspPheValGluIleHisGlyLysHisGlnArgGlnAspAspHisGlyTyrIle 60  
 DB 283 GAAGATTTCTGGAGATCCATGGCAACACACACAGAGAGCAGGATGACCTGGCTACAT 342  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 DB 343 TCCCGTGAATTCACCGTCGTACCGTCTGCCCTTCCATGTGGACAGTCCGCCCTCTCC 402  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 DB 403 TGCTCTGTCTGGAGTGGATGCTGACCTTCTTGCCCCCAAGTCCAGTCTGGCTG 462  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 DB 463 GATGTGGCCACAGCAGAGGGCCATTCCCGTGTCCAGGGAGGAGAGCCAGCTCGGCA 522  
 QY 121 ProSerSer 123  
 DB 523 CCCTCGTCC 531

RESULT 9  
 BOVCRYA  
 LOCUS BOVCRYA 793 bp mRNA linear MAM 26-APR-1993  
 DEFINITION Bovine alpha-A-crystallin gene, complete cds.  
 ACCESSION M26142

M26142.1 GI:162909  
 crystallin.  
 SOURCE Bos taurus (cow)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovinae; Bos.  
**REFERENCE** 1 (bases 1 to 793)  
**AUTHORS** Hay, R.E. and Petrash, J.M.  
**TITLE** Nucleotide sequence of a bovine lens alpha A-crystallin cDNA  
**JOURNAL** Biochem. Biophys. Res. Commun. 148 (1), 31-37 (1987)  
**MEDLINE** 88049675  
**PUBMED** 3675580  
**COMMENT** Original source text: Bovine lens, cDNA to mRNA, clone  
 pBL-alpha-A2-1.  
**FEATURES**  
 Location/Qualifiers  
 source  
 1..793  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 31..552  
 /note="alpha-A-crystallin"  
 /codon\_start=1  
 /protein\_id="AAA30471.1"  
 /db\_xref="GI:162910"  
 /translation="MDIAIQHPWFKRLGPVPSRLFOQFGEGLFEYDLPLFLSSTI  
 SPYRQSLFRVLDSGISVSRDRDKFVFLDVKHFSPELITVKVQEDFVEIHGKNE  
 RQDDHGYSIRSEFHRRYRLPSNVDSQALSCLSDGMLTFSKPKIPSGVDAGHSRAIP  
 VSREKFPSSAPSS"  
**ORIGIN** 29 bp upstream of NcoI site.

Alignment Scores:  
 Pred. No.: 2,786-58 Length: 793  
 Score: 598.00 Matches: 115  
 Percent Similarity: 97.56% Conservative: 5  
 Best Local Similarity: 93.50% Mismatches: 3  
 Query Match: 93.58% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x BOVCRYA (1-793)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 DB 181 TCTCTTCCGACAGTGTGGACTCCGGCATCTCTGAGTCCGATCTGACCGGACAG 240  
 QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
 DB 241 TTTGTCTCTTCTGGATGTGAAGCACCTCTCTCCGAGGACCTCACCGTGAAGTGTGAG 300  
 QY 41 AspAspPheValGluIleHisGlyLysHisGlnArgGlnAspAspHisGlyTyrIle 60  
 DB 301 GAGGACTTCTGGAGATCCAGCGCAAGCACACAGAGCGGAGATGACCTGGCTACATC 360  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 DB 361 TCCCGGAGTTCCACCGCTGCTGCTCCAACTGAGGACAGTCCGACCTCTCC 420  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 DB 421 TGCTCTGTCTGGAGTGGATGCTGACCTTCTTGCCCCCAAGTCCATCTGGCGTG 480  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 DB 481 GATGTGGCCACAGCAGAGGGCCATTCCCGTGTCCAGGGAGGAGAGCCAGCTCTGCG 540  
 QY 121 ProSerSer 123  
 DB 541 CCCTCGTCC 549

RESULT 10  
 RNACRY  
 LOCUS RNACRY 885 bp mRNA linear ROD 08-JAN-1999  
 DEFINITION Partial sequence of messenger RNA for rat alpha-a2-crystallin.

ACCESSION V01219 J00715  
 VERSION V01219.1 GI:55598  
 KEYWORDS complementary DNA; crystallin.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 885)  
 AUTHORS Moormann, R.J., van der Velden, H.M., Dodemont, H.J., Andreoli, P.M.,  
 Bloemendal, H., and Schoenmakers, J.G.  
 TITLE An unusually long non-coding region in rat lens alpha-crystallin  
 messenger RNA  
 JOURNAL Nucleic Acids Res. 9 (19), 4813-4822 (1981)  
 MEDLINE 82081811  
 PUBMED 6171772  
 FEATURES  
 source  
 1..885 Location/Qualifiers  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 <1..885  
 <1..368  
 /codon\_start=3  
 /product="crystallin"  
 /protein\_id="CAA24530.1"  
 /db\_xref="GI:809074"  
 /translation="PRTVLDGISEVRSRDQKVFILDFVLDVHKFSPEDLTVKVLDFVEI  
 HGKNERQDDHGYSIRFHRRLPSNVDSALSCSLADGLMTFPGPKVQSGLDGHS  
 SERAIPVSRKEKPSAPSS"  
 885  
 polyA\_site  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,978-57 Length: 885  
 Score: 591.00 Matches: 113  
 Percent Similarity: 97.52% Conservative: 5  
 Best Local Similarity: 93.33% Mismatches: 3  
 Query Match: 92.43% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x RNACRY (1-885)  
 QY 3 PheArgThrValLeuAspSerGlyIleSerGluValArgSerAspAspLysPheVal 22  
 DB 3 TTCCGACAGTGTGGACTCCGGCATCTCTGAGGTCGATCTGACCGGACAGATTTC 62  
 QY 23 IlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAsp 42  
 DB 63 ATCTTCTTGGATGTGAAGCACCTCTCTCTGAGGACCTCACCGTGAAGTACTGGAAGAT 122  
 QY 43 PheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArg 62  
 DB 123 TTCGTGGAGATCCATGGCAACACACAGAGGAGGATGACCATGGCTTACATTTCCCGT 182  
 QY 63 GluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSer 82  
 DB 183 GAATTTACCGTCGCTACCGTCTGCTTCCATGTGGACAGTCCGCCCTCTCTGCTCC 242  
 QY 83 LeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAla 102  
 DB 243 TTGCTGGGATGGCATGTGACCTTCTCTGGCCCAAGGTCAGTCTGGCTTGGATGCT 302  
 QY 103 ThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAlaProSer 122  
 DB 303 GCCACAGCAGAGGAGGCCATTCCCGTGTACGGGAGGAGGCCAGCTCGGACCCCTCG 362  
 QY 123 Ser 123  
 DB 363 TCC 365  
 RESULT 11  
 MMU310308

LOCUS MMU310308 796 bp mRNA linear ROD 03-APR-2001  
 DEFINITION Mus musculus mRNA for alpha-A-crystallin (Cryaa gene).  
 ACCESSION AJ310308  
 VERSION AJ310308.1 GI:13548627  
 KEYWORDS alpha-A-crystallin; cryaa gene.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 AUTHORS Graw, J.  
 TITLE Characterization of a new, dominant V124E mutation in the mouse  
 alpha-A-crystallin encoding gene  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 796)  
 AUTHORS Graw, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-APR-2001) Graw J., Institute of Mammalian Genetics,  
 GSF-National Research Center for Environment and Health,  
 Ingolstaedter Landstr. 1, Neuherberg, Germany, D-85764, GERMANY  
 FEATURES  
 source  
 1..796 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C3H/E1"  
 /db\_xref="taxon:10090"  
 /tissue\_type="Lens"  
 1..796  
 /gene="Cryaa"  
 52..642  
 /gene="Cryaa"  
 /function="lens structural protein; chaperone"  
 /note="alternative splicing"  
 /codon\_start=1  
 /product="alpha-A-crystallin"  
 /protein\_id="CAC35974.1"  
 /db\_xref="GI:13548628"  
 /translation="MDVTIOHPWPKRALGPPYPSRLFDORFEGFLFEYDLLPFLSSTI  
 SPYRQSLFRVLDSGISELMTWFMVHOPHAGNPKNPVKVRSRDRDKFVFLDVKH  
 FSPEDLTVKVLDFVEIHGKNERQDDHGYSIRFHRRLPSNVDSALSCSLADG  
 MLTSPGPKVQSGLDGHSERAIPIVSRKEKPSAPSS"  
 241..309  
 /gene="Cryaa"  
 /note="alternative splicing"  
 exon  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,918-56 Length: 796  
 Score: 577.50 Matches: 115  
 Percent Similarity: 82.19% Conservative: 5  
 Best Local Similarity: 78.77% Mismatches: 3  
 Query Match: 90.38% Indels: 23  
 DB: 10 Gaps: 1  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x MMU310308 (1-796)  
 QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGlu-----ValArgSerAsp 17  
 DB 202 TCCCTCTCCGACTGTGCTGGACTCGGGCATCTCTGAGCTCATGACCCATATGTGGTTT 261  
 QY 14 -----ValArgSerAsp 17  
 DB 262 GTAATGCACCAACACATGCTGGAAACCCCAAGAACACCCCGTCAAGTCCGATCTGAC 321  
 QY 18 ArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrVal 37  
 DB 322 CGGACAAAGTTGTTCATCTCTTGGACGTGAAGCATTCTCTCTGAGGACCTCACCGTG 381  
 QY 38 LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHis 57  
 DB 382 AAGTACTGGAGGATTTGTGGAGATTACCGCAACACACAGAGAGCAGATGACCAT 441  
 QY 58 GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSer 77



```

Alignment Scores:
Pred. No.: 3,97e-53 Length: 447
Score: 550.00 Matches: 103
Percent Similarity: 99.09% Conservative: 6
Best Local Similarity: 93.64% Mismatches: 1
Query Match: 86.07% Indels: 0
DB: 4 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x AJ617725 (1-447)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 118 TCCTCTTCGCGACCGTCTCGACTCCGGCATCTCCGAGTCCGGTCTGATCGGGACCA 177
QY 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40
Db 178 TTCCTCATCTCTGACGCTGAGCACTTCTCCCTGAGGACCTGACTGTGAAGGTGCAG 237
QY 41 AspAspPheValGluIleHisGlyIleHisAsnGluArgGlnAspAspHisGlyIle 60
Db 238 GATGACTTTGTGGAGATCCATGGCAACCAATGAGAGCGAGGACGACCAAGCTACATC 297
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 298 TCCCGTGAGTTCCACCGCGCTACCGCTGCCCTCCACGCTGAGCAGTCTGCACTCTCT 357
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 358 TGCTCCCTCTCTCGGACCGGATGCTGACTTCTGTGGCCCCCAAGATCCAGTCTGGCATG 417
QY 101 AspAlaThrHisAlaGluArgAlaIlePro 110
Db 418 GATGCCAGCCACAGTGGAGAGCCATCCCC 447

RESULT 14
AJ617726 477 bp mRNA linear VRT 15-JUL-2004
LOCUS Sphenodon punctatus partial mRNA for alphaA-crystallin (cryaa
DEFINITION
ACCESSION AJ617726
VERSION AJ617726.1 GI:50344348
KEYWORDS alphaA-crystallin; cryaa gene.
SOURCE Sphenodon punctatus (tuatara)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
AUTHORS van Rheede, T. and de Jong, W.W.
TITLE The alpha-crystallins of the platypus Ornithorhynchus anatinus:
JOURNAL Origin of the alternatively spliced exon alphaAins and implications
REFERENCE for mammalian phylogeny
AUTHORS Unpublished
JOURNAL (bases 1 to 477)
TITLE Direct Submission
JOURNAL Franck, E.
FEATURES Submitted (15-DEC-2003) Franck E., Biochemistry, Radboud University
REFERENCE Nijmegen, 9101, 6500 HB Nijmegen, NETHERLANDS
AUTHORS Location/Qualifiers
JOURNAL 1. 477
SOURCE /organism="Sphenodon punctatus"
/locus="AJ617726"
/db_xref="taxon:8508"
1. 477
1. 477
<1. 477
/gene="cryaa"
/codon_start=1
/product="alphaA-crystallin"
/db_xref="GI:50344349"
/translat="PLIPSRFLDFGGEGLPEYDILLPLFGSTIPYRQSFRTVLRS
GISEVSRDRDKFTIFLDVKEFSPEDLSVKIIDDVEIFGKHNERQVDHGVISREHFR
YRLPSNVDSQSAITCSLSDGCMWTSPAKPVQSNMDPSHSRPIVSRREKTSAPSS"

gene
CDS

US-10-657-740-1_COPY_51_173 (1-123) x AJ617726 (1-477)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 106 TCCTTTCCGCGACCGTCTCGGAATCAGGCATTTTCAGAGGTGAGATCTGACCGGACAAG 165
QY 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40
Db 166 TTTCAATCTTCTCGGATGTAAACACTTCTCTCTGAGAGATTGAGTGTGAAGATTATT 225
QY 41 AspAspPheValGluIleHisGlyIleHisAsnGluArgGlnAspAspHisGlyIle 60
Db 226 GATGACTTTGTGGAATCCACGCAAGCACAAATGAGACAGGTAGACCATGGCTACATC 285
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 286 TCCCGGCAATTTCCACCGGAGGTACCGCTGCCCTTCCAAATGTGGACCAATCCGCCATCACT 345
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 346 TGCTCCCTCTCTGTGTGAGCGCATGATGACCTTCTCTGCCCCCAAGGTCCAGTCTAACATG 405
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 406 GACCCGACCCACGAGAGAGACCCATTCCTGTATCCCGTGAAGAGAGCCACCTCGCGCG 465
QY 121 ProSerSer 123
Db 466 CCTTCTCTCC 474

RESULT 15
RCAACRYST 522 bp mRNA linear VRT 18-SEP-1995
LOCUS R.catesbeiana mRNA for alpha-crystallin alpha A subunit.
DEFINITION
ACCESSION X85205
VERSION X85205.1 GI:732915
KEYWORDS aa-crystallin.
SOURCE Rana catesbeiana (bullfrog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana;
AUTHORS Aquarana.
TITLE
JOURNAL
MEDLINE
PUBMED 95283564
REFERENCE 2 (bases 1 to 522)
AUTHORS Chiou, S.H.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-1995) S.H. Chiou, Institute of Biological
REFERENCE Chemistry, Academia Sinica, Institute of Biochemical Sciences,
AUTHORS National Taiwan University, P.O. Box 23-106, Taipei 10098, TAIWAN
JOURNAL Sequence overlapping with that under the acc#X00716.
FEATURES Location/Qualifiers
SOURCE 1. 522
/mol_type="mRNA"
/db_xref="taxon:8400"
1. 522
CDS

```

```

/codon_start=1
/product="alpha-crystallin alpha A subunit"
/protein_id="CAA59471.1"
/db_xref="GI:732916"
/db_xref="GOA:Q91311"
/db_xref="UniProt/Swiss-Prot:Q91311"
/translation="MDIAGHPWKRAGLPYNNRLDFQVGEGMFYDILFPFLSTV
SPYRHSLPFGMSDGI SEVARSDFRTNFIDVHPGDOLTKVILDFDYIIRKHS
RODDGYIGIREFHRYRLPSNLDDQSSITCSAGIULTFSPGKMWGLSDSHSERDIP
VSEKEXPTSAAPS"

```

## ORIGIN

Alignment Scores:	
Pred. No.:	2,286-52
Score:	544.00
Percent Similarity:	93.50%
Best Local Similarity:	82.11%
Query Match:	85.13%
DB:	5
Length:	522
Matches:	101
Conservative:	14
Mismatches:	8
Indels:	0
Gaps:	0

US-10-657-740-1 COPY 51 173 (1-123) x RCAACRYST (1-522)

QY	1	SerLeuPheArgThrValLeuAspSerClyIleSerGluValArgSerAspArgAspLys	20
Db	151	AGCCTCTCCGAGGATTCAAGATCCGGCATTTTCAGGTGGCTTCAGATCCGATCGT	210
QY	21	PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValLysValGln	40
Db	211	TTCAACATTACTTGGATGTGAGCATTTCTACCCGATGATCTGACTGTGAAGATTCTG	270
QY	41	AspAspPheValGluIleHisClyIleHisAsnGluArgGlnAspHisGlyTyrlle	60
Db	271	GATGACTTGTGGAAATCCATGGAAAGCACAGCAGGAGGAGGATGACCATGCTACATT	330
QY	61	SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer	80
Db	331	TCCCGGGAGTTCCACCGCGCTACCGCTCCCTCCAACTGGACCAATCTCTATCAGC	390
QY	81	CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProValIleGlnThrGlyLeu	100
Db	391	TGCTCACTCTCTGTGATGGCATTTGACCTTCAGCGGACCAAAATGATGTGAGCCGCTG	450
QY	101	AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla	120
Db	451	GACTCCAGCCACGAGGAGGCGCTATCCCTGTGTGCAGAGAGAAAGCCACCTCTGCC	510
QY	121	ProSerSer	123
Db	511	CCCTCCTCC	519

RESULT 16	
AJ617724	
LOCUS	558 bp mRNA linear MAM 15-JUL-2004
DEFINITION	Ornithorhynchus anatinus partial mRNA for alphaA-crystallin (cryAA gene).
ACCESSION	AJ617724
VERSION	AJ617724.1 GI:50344344
KEYWORDS	alphaA-crystallin; cryAA gene.
SOURCE	Ornithorhynchus anatinus (platypus)
ORGANISM	Ornithorhynchus anatinus

REFERENCE	1	van Rheede, T. and de Jong, W.W.
AUTHORS		
TITLE		The alpha-crystallins of the platypus <i>Ornithorhynchus anatinus</i> : Origin of the alternatively spliced exon alphaAins and implications for mammalian phylogeny
JOURNAL		Unpublished
REFERENCE	2	Franch, E.
AUTHORS		
TITLE		Direct Submission
JOURNAL		Submitted (15-DEC-2003) Franch E., Biochemistry, Radboud University Nijmegen, 9101, 6500 HB Nijmegen, NETHERLANDS

```

FEATURES
source
Location/Qualifiers
1..558
    /organism="Ornithorhynchus anatinus"
    /mol_type="rRNA"
    /db_xref="taxon:9258"
1..558
    /gene="cryaA"
<1..558
    /gene="cryaA"
    /codon_start=1
    /product="alphaA-cryatallin"
    /protein_id="CAF02100.1"
    /db_xref="GI:50344345"
    /translation="RALGPLYPSRLFDQPFGEGLFVE  
VLESGISLMHVMFVHKSHAGNPKNPAKVRSDRI  
DEFVIGHKHSERODHGYISRETHRYRLPSNDVQ  
NLDSSHSDPSIPISREEVPTSPSS"
157..225
    /gene="cryaA"
    /note="ins exon"
exon

```

## ORIGIN

Alignment Scores:	
Pred. NO.:	1.42e-50
Score:	528.50
Length:	558
Matches:	102
Percent Similarity:	80.14%
Best Local Similarity:	Conservative: 15
Query Match:	Mismatches: 6
DB:	Indels: 23
	Gaps: 1
	4

US-10-657-740-1 COPY 51 173 (1-123) x AJ617724 (1-558)

```

1 SerLeuPheArgThrValLeuAspSerGlyIleSerGlu-----13
| | | | | | | | | | | | | | | | | | | | |
118 TCTCTGTTCCGCACGGCTCTGGAGTCGGGCAATTCGAGCTCATGCCCATGTATGCTTT177
| | | | | | | | | | | | | | | | | | | | |
14 -----ValArgSerAsp17
| | | | | | | | | | | | | | | | | | | | |

```

QY	18	ArgAspLysPheValIlePheLeuaspValLysHisPheSerProGluaspLeuThrVal	37
Db	178	GTAATGCACAAATACATCTGCTGGAAATCCCAAGAA	237
Db		CAACCTCCCAAGGTGGCATCTGAC	238
Db		CGGACAAGTTTGTTCATCTCTGGATGAAGCAATTTCTCCCAAGATCTGACGGTG	297

**Qy**    38 LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHis 57  
       ||||| : : : : :  
**Db**    298 AAGGTGGTGAATGAGTTTGCGAGATTTCATGCCAAGCACAGTGGACAGACAGGATGCACC 357  
                                :

Qy 58 GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSer 77  
 358 GTTACATTTCCGGGGAAATTCATCGCCGGTACGCGCTTGCCTTCAACGTGGACAGGCC 417  
 Db

Qy	78	AlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGln	97
		: : : :	
Db	418	TCAGTCTCCTGTCCCTGCTCTGTATGGCATGTTGACCTTCTCTGGCTCCAAGGTCCAA	477

Qy 98 ThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro 117  
::: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 478 TCCAACCTGGATTCCAGCCACAGTGACCGCATCCCATCCCTATCTCCCGGAGGAAGTGCCC 537

Qy	118	Thr	Ser	Ala	Pro	Ser	Ser	123
Db	538	ACCTCAAC	CTCCCTCT	CTCT	CTCT	CTCT	CTCT	555

## RESULT 17

LOCUS	RTCRYA	765 bp	mRNA	linear	VRT 06-JUL-1989			
DEFINITION.	Frog mRNA fragment for alpha-A2-crystallin.							
ACCESSION	X00716	X00058						
VERSION	X00716.1	GI:64292						
KEYWORDS	alpha-crystallin; crystallin.							
SOURCE	Rana temporaria (common frog)							
ORGANISM	Rana temporaria							



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana;  
Rana.

## REFERENCE

1 Tomarev, S.I., Zinovleva, R.D., Dolgilevich, S.M., Kravay, A.S.,  
Skraybin, K.G. and Gause, G.G. Jr.

The absence of the long 3'-non-translated region in mRNA coding for  
eye lens alpha A2-crystallin of the frog (*Rana temporaria*)

## JOURNAL

84005173

## MEDLINE

6604666

## PUBMED

2 (bases 1 to 765)

## REFERENCE

Tomarev, S.I., Zinov'eva, R.D., Kraev, A.S., Skryabin, K.G. and  
Gause, G.G.  
Primary structure of cloned cDNA coding alpha-A2 crystallin of the  
eye lens of the frog *Rana temporaria*  
Dokl. Biochem. 271, 277-280 (1984)  
Data kindly reviewed (05-11-1985) by S.I. Tomarev  
Sequence 1 to 185 is complementary to sequence 553 to 369 and  
probably a cloning artefact; amino terminal amino acids are  
missing.

## FEATURES

source Location/Qualifiers

1..765 /organism="Rana temporaria"

/mol\_type="mRNA"

/db\_xref="taxon:8407"

## misc\_feature

1..185

/note="artefactual sequence"

## CDS

185..632

/note="unnamed protein product; Protein sequence is in

conflict with the conceptual translation; alpha-A2

crystalline (aa 25 to 173)"

/codon\_start=1

/protein\_id="CAA25308.1"

/db\_xref="GI:1334774"

/db\_xref="GOA:P02508"

/db\_xref="UniProt/Swiss-Prot:P02508"

/translation="QVFGEGMFDYDLFPFLTSTVSPHYRHGILRGFMDSGISEVRSR

DRFTNLNLYKHFSPDILTKILDDVEIHGHSEKQDDHGVISREHFHYRLPSNLDQ

SSISGSLSDAGILTFSGPKMNSLVSHSERPIVSRREKETSAPSS"

## misc\_feature

744..749

/note="polyadenylation signal"

## polyA\_site

765

/note="polyadenylation site"

## ORIGIN

Alignment Scores:

Pred. No.: 2 45e-49 Length: 765

Score: 519.00 Matches: 97

Percent Similarity: 90.98% Conservative: 14

Best Local Similarity: 79.51% Mismatches: 11

Query Match: 81.22% Indels: 0

DB: 5 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x RTCYA (1-765)

QY 2 LeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLysPhe 21

Db 267 TCCTCCCGAGGATTCATGCATCCCGCATCTCTGAGTGCGTTCAGATCGCATCGTTC 326

QY 22 ValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGlnAsp 41

Db 327 ACCATTACTTGGACGTGAAGATTCTCCCGCATGATGATGATGATGATGATGATGAT 386

QY 42 AspPheValGluIleHisGlyIleHisGlnAsnGluArgGlnAspAspHisGlyTyrIleSer 61

Db 387 GACTTGTGGAAATCATGGAAGACACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 446

QY 62 ArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCys 81

Db 447 CGGAGTGTCCACCGCGCTACCGCTCCCTCCCAACCTGGACCAATCTCTATCAGCTGC 506

QY 82 SerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAsp 101

Db 507 TCACCTCTCGCGATTCGACCTTACGCGACCAAAATGATGTCAAACCTGTC 566

QY 102 AlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAlaPro 121

Db 567 TCACGCCACAGCAGGAGGCCATCCCTGTGTCCAGAGAGGAAAGCCACCTTGCCTCC 626

QY 122 SerSer 123

Db 627 TCCTCC 632

## RESULT 18

D86299

LOCUS

D86299

DEFINITION

Cynops pyrrhogaster mRNA for newt alpha A-crystallin, complete cds.

ACCESSION

D86299

VERSION

D86299.1

GI:19168451

KEYWORDS

Cynops pyrrhogaster (Japanese firebelly newt)

Cynops pyrrhogaster

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;

Cynops.

REFERENCE

1 Mizuno, N., Agata, K., Sawada, K., Mochii, M. and Eguchi, G.

Expression of crystallin genes in embryonic and regenerating newt

lenses

Dev. Growth Differ. 44 (3), 251-256 (2002)

JOURNAL

22054970

MEDLINE

12060074

PUBMED

REFERENCE

2 (bases 1 to 1361)

AUTHORS

Mizuno, N.

TITLE

Direct Submission

Submitted (27-JUN-1996) Nobuhiko Mizuno, Osaka University,

Institute for Molecular and Cellular Biology, Kondoh

Differentiation Signaling Project ERATO; 1-3 Yamadaoka, Suita-shi,

Osaka 565-0871, Japan (E-mail: nmizuno@imcb.osaka-u.ac.jp,

tel:06-6879-7964, fax:06-6877-1738)

FEATURES

Location/Qualifiers

1..1361

/organism="Cynops pyrrhogaster"

/mol\_type="mRNA"

/db\_xref="taxon:8330"

142..663

/codon\_start=1

/product="newt alpha A-crystallin"

/protein\_id="BAB85811.1"

/db\_xref="GI:19168452"

/translation="WDITIQHFWKRALGPPYGRILFDQFFGDLFDVLEFPFLSSTV

SPYQSGMFRNYLDSGISEVSRGRDKPQIYLDVHKHSPEDLSVKILDDYVEIHGHSD

RQDDHGYVSRFHRRLPASPVDQSSITCSLSLTDGMLTFSGTGMSLSSSHGRPIP

VSREKPKPAAQSS"

ORIGIN

Alignment Scores:

Pred. No.: 1 11e-47 Length: 1361

Score: 507.00 Matches: 93

Percent Similarity: 89.43% Conservative: 17

Best Local Similarity: 75.61% Mismatches: 13

Query Match: 79.34% Indels: 0

DB: 5 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x D86299 (1-1361)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20

Db 292 TCATGTTCCGCACTACCTGACCTCTGCATATCCGAGGTGAGGTCTGCGCGCACAG 351

QY 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40

Db 352 TTCAGATTACCTGGACGTCAAGCACTCTCACCGGAAGACCTAAGTGTCAAGATCTC 411

QY 41 AspAspPheValGluIleHisGlyIleHisGlnAsnGluArgGlnAspAspHisGlyTyrIle 60



```

Alignment Scores:
Pred. No.: 1,11e-47 Length: 448
Score: 502.00 Matches: 92
Percent Similarity: 93.64% Conservative: 11
Best Local Similarity: 93.64% Mismatches: 7
Query Match: 78.56% Indels: 0
DB: 5 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x TSU31938 (1-448)

QY 1 SerLeuPheArgThrValLeuAspSerGlyHisValArgSerGluValArgSerAspArgAspLys 20
Db 119 TCCTCTTCGCGACCGTCTCGGAATTCAGGATTCAGAGGTGAGGTCTGACCGGACAAAG 178
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40
Db 179 TTTTACAATCTCTGGATGTAAACACTTCTCCCGAAGATCTGAGTGTGAAGATTATG 238
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60
Db 239 GATGACTTTGTGAAATTCATGTCGAAGCACAATGAGACAGACGACCATGGCTACATT 298
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 299 TCCCGGAATTCACCGCAGATACCGCTGCTTCCACGTCGACCAATCTGCCATCAC 358
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 359 TGCTCCTGCTGCTGATGCGATGCTGACTTCTCTGGCCCAAAAGTCCAGTCCCAACATG 418
QY 101 AspAlaThrHisAlaGluArgAlaIlePro 110
Db 419 GACACCAGTATAGCGAGAGCCCATTCCT 448

RESULT 21
D88185 LOCUS
DEFINITION Xenopus laevis mRNA for alpha A crystallin, complete cds.
ACCESSION D88185
VERSION D88185.1 GI:4589827
KEYWORDS alpha A crystallin.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
Mizuno,N., Mochii,M., Takahashi,T.C., Ezuchi,G. and Okada,T.S.
Lens regeneration in Xenopus is not a mere repeat of lens
development, with respect to crystallin gene expression
Differentiation 64 (3), 143-149 (1999)
99251119
PUBMED 10234811
2 (bases 1 to 705)
Mizuno,N.
Direct Submission
Submitted (30-SEP-1996) Nobuhiko Mizuno, Biohistory Research Hall,
Laboratory Div.; Murasaki Chou 1-1, Takatuki, Osaka 569-11, Japan
(E-mail:Nobuhiko.Mizuno@ate.brh.co.jp, Tel:0726-81-9751,
Fax:0726-81-9757)
Sequence updated (12-Apr-1999).
COMMENT Location/Qualifiers
FEATURES
source
1..705
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
67..582
/codon_start=1
/product="alpha A crystallin"
/protein_id="BAA76897.1"
/db_xref="GI:4589828"
/translation="MDITIQHPWPKRSLGPPFNRLFDQVFGGMFDFLFPWMSSTI
SPYTKQNLRSGLDGSIVSRDRDFVNLNVKHFSPEDLSVKVHDDFVEIHKINE

CDS

```

```

RQDHGYISRFHRRYRLPSNMDSVSCSLSADGILTLFGPKLQSNMDSHSDRTIP
VSKEKSGSSS"
705
/Note="22 a nucleotides"

polyA_site
ORIGIN
Alignment Scores:
Pred. No.: 2,58e-46 Length: 705
Score: 492.00 Matches: 91
Percent Similarity: 90.83% Conservative: 18
Best Local Similarity: 75.93% Mismatches: 11
Query Match: 77.00% Indels: 0
DB: 5 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x D88185 (1-705)

QY 1 SerLeuPheArgThrValLeuAspSerGlyHisValArgSerAspArgAspLys 20
Db 217 AACCTCTCCAGAGGATACCTGGACTCTGGGATCTCTGAGGTGCGTTGAGACCGTGACCGT 276
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40
Db 277 TTTGTCAATTACCTGGATGTGAAGCATTTCTCCCTGAAGATCTGAGCGTCAAAAGTCCAT 336
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60
Db 337 GATGACTTTGTAGAGATTTCATGGGAAACACACAGCAGCAGGATGATCAGGATACATC 396
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 397 TCCCGAATTCATCGCCGATATCGTCTTCTTCCATATGATGATCAGAACTCTGTGAGC 456
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 457 TGCACTCTGTCTCGGACGGATCTCTTCTTTCGTCCTCCAACTGCAATCCACATG 516
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 517 GACTCCAGCCACAGGATAGGACCATCTCTGTCTCCAGGAGAGAAATCAGGCTCATCC 576

RESULT 22
EUDLAACRYS LOCUS
DEFINITION Eudromia elegans alpha A-crystallin mRNA, partial cds.
ACCESSION L25850
VERSION L25850.1 GI:538348
KEYWORDS alpha A-crystallin.
SOURCE Eudromia elegans (elegant crested-tinamou)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Palaeognathae; Tinamiformes; Tinamidae;
Eudromia.
1 (bases 1 to 448)
Caspers,G.J., Wattel,J. and de Jong,W.W.
Alpha A-crystallin sequences group tinamou with ratites
Mol. Biol. Evol. 11 (4), 711-713 (1994)
94359400
PUBMED 8078410
COMMENT Original source text: Eudromia elegans eye lens cDNA to mRNA.
FEATURES
Location/Qualifiers
source
1..448
/organism="Eudromia elegans"
/mol_type="mRNA"
/db_xref="taxon:8805"
/tissue_type="eye lens"
<1..448
/codon_start=2
/product="alpha A-crystallin"
/protein_id="AAA49254.1"
/db_xref="GI:538349"
/translation="PALGPLPSRLFDQFGGLLEYDLLPLFSTISPYTQSLFRS
VLESLGSEVRSDREKFTMLDVKHFSPEDLSVKIIDDVFEIHKHSEKQDDHGYSRE
FHRRYRLPSNMDSVSCSLSDGMLTFSGPKVQANMDPSHSRPIP"

CDS

```

```

ORIGIN
Alignment Scores:
Pred. No.: 5.6e-46 Length: 448
Score: 487.00 Matches: 89
Percent Similarity: 92.73% Conservatives: 13
Best Local Similarity: 80.91% Mismatches: 8
Query Match: 76.21% Indels: 0
Db: 5 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x EUDLAACRYS (1-448)
QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 119 TCTCTCTCCGAGCGTCTGGAGTCAGCATTTTCAGAGGTGAGATCTGACCGGGAATAA 178
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40
Db 179 TTTACAATCATGCTGGATGTAACACATCTCTCTGAAGACCTAAGCGTGAAGATCAT 238
QY 41 AspAspPheValGluIleHisGlyLysHisGlnArgGlnAspAspHisGlyTyrIle 60
Db 239 GATGACTTTGTGGAATCCATGCGAAGCACAGTGAAGACAGACGACCATGCTGATCATC 298
QY 61 SerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 299 TCCCGTGAAATTCACCGCGGTACCGCTCCACCGTGACGACCATGCTGCCATCACC 358
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 359 TGCTCCCTCTCCAGTACGCGATGCTGACCTTCTCCGCGCCCAAGTCCAGCGCAACATG 418
QY 101 AspAlaThrHisAlaGluArgAlaIlePro 110
Db 419 GACCCAGCCACAGCAGAGACCCATCCCT 448

RESULT 23
APRAAC APRAAC 448 bp mRNA linear VRT 09-SEP-2004
LOCUS A.platyrrhynchos mRNA for alpha-A-crystallin.
ACCESSION X96592
VERSION X96592.1 GI:1945629
KEYWORDS alpha-A-crystallin.
SOURCE Anas platyrhynchos
ORGANISM Anas platyrhynchos
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
REFERENCE 1
AUTHORS Caspers,G.J., Uit de Weerd,D., Wattel,J. and de Jong,W.W.
TITLE alpha-Crystallin sequences support a galliform/anseriform clade
JOURNAL Mol. Phylogenet. Evol. 7 (2), 185-188 (1997)
MEDLINE 97271643
PUBMED 9126559
REFERENCE 2 (bases 1 to 448)
AUTHORS Caspers,G.J.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1996) G.J. Caspers, University of Nijmegen,
Department of Biochemistry, P.O.Box 9101, NL- 6500 HB Nijmegen,
NETHERLANDS
FEATURES
source
1. 448
/organism="Anas platyrhynchos"
/mol_type="mRNA"
/strain="Peking"
/db_xref="taxon:8839"
/tissue_type="lens"
<1. .>448
/codon_start=2
/product="alpha-A-crystallin"
/protein_id="CAA65410.1"
/db_xref="GI:1945630"
/db_xref="GOA:O12984"
/db_xref="UniProt/Swiss-Prot:O12984"
CDS

ORIGIN
Alignment Scores:
Pred. No.: 7.27e-46 Length: 448
Score: 486.00 Matches: 89
Percent Similarity: 91.82% Conservatives: 12
Best Local Similarity: 80.91% Mismatches: 9
Query Match: 76.06% Indels: 0
Db: 5 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x APRAAC (1-448)
QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 119 TCCCTCTCCGAGCGTCTGGAGTCGGCATTTTCAGAGGTGAGTCTGACCGGCAACAG 178
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40
Db 179 TTTACGATCATGCTGGATGTAACACATCTCTCTGAAGATCTGAGCGTGAAGATATC 238
QY 41 AspAspPheValGluIleHisGlyLysHisGlnArgGlnAspAspHisGlyTyrIle 60
Db 239 GATGACTTTGTGGAATCCATGCGAAGCACAGTGAAGACAGACGACCATGCTGATCATC 298
QY 61 SerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 299 TCCCGTGAGTTTCCCGCGGTACCGCTCCACCGTGACGACCATGCTGCCATCACC 358
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 359 TGCTCCCTCTCCGCGCGCATGCTGACCTTCTCAGCGCCCAAGTCCCGTCCCAACATG 418
QY 101 AspAlaThrHisAlaGluArgAlaIlePro 110
Db 419 GACCCAGCCACAGCAGAGACCCATCCCT 448

RESULT 24
APRAAC APRAAC 448 bp mRNA linear VRT 09-SEP-2004
LOCUS C.livia mRNA for alpha-A-crystallin.
ACCESSION X96593
VERSION X96593.1 GI:1945730
KEYWORDS alpha-A-crystallin.
SOURCE Columba livia (domestic pigeon)
ORGANISM Columba livia
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
REFERENCE 1
AUTHORS Caspers,G.J., Uit de Weerd,D., Wattel,J. and de Jong,W.W.
TITLE alpha-Crystallin sequences support a galliform/anseriform clade
JOURNAL Mol. Phylogenet. Evol. 7 (2), 185-188 (1997)
MEDLINE 97271643
PUBMED 9126559
REFERENCE 2 (bases 1 to 448)
AUTHORS Caspers,G.J.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1996) G.J. Caspers, University of Nijmegen,
Department of Biochemistry, P.O.Box 9101, NL- 6500 HB Nijmegen,
NETHERLANDS
FEATURES
Location/Qualifiers
source
1. 448
/organism="Columba livia"
/mol_type="mRNA"
/db_xref="taxon:8932"
/tissue_type="lens"
<1. .>448
/codon_start=2
/product="alpha-A-crystallin"
/protein_id="CAA65411.1"
/db_xref="GI:1945731"
CDS

```

```

/db_xref="GOA:O12988"
/db_xref="UniProt/Swiss-Prot:O12988"
/translation="RALGPIPLPSRLPDQFFGGLLEVDLLPWFSSITSPYRSLFRS
VLESGISVRSRDEKFTMLDVKHSPEDLSVLKIDDFVEIHGHSERQDDHGYISRE
PHRYRLPANVDQAATCSLNDGMLTFSGPKVPANMDASHGERPIP"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 9.92e-45 Length: 448
Score: 476.00 Matches: 87
Percent Similarity: 90.91% Conservative: 13
Best Local Similarity: 79.09% Mismatches: 10
Query Match: 74.49% Indels: 0
DB: 5 Gaps: 0

```

US-10-657-740-1\_COPY\_51\_173 (1-123) x CLRAAC (1-448)

```

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspAspLys 20
Db 119 TCCCTCTTCGCGAGTGTGGAGTCAGGATTCAGAGTGAGGTCTGACCGGAAAG 178
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40
Db 179 TTTACAATCATGTGGATGTAAACACTTCTTCCGAAGACCTGAGTGTGAAGATTAT 238
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60
Db 239 GAGCACTTTGTGAAATCCATGTCAGACAGCAGTGAAGACAGCAGCACCGGTTACATC 298
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 299 TCCCGGAAATTCACCGCGGTACCGCCCTGCGCGCCCAACGTCGACCGCTCCATCACC 358
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 359 TGTCTCCTGTCCAAACAGCGCATGCTGACCTTCTCGGGCCCCCAAGTCCCGCCCAACATG 418
QY 101 AspAlaThrHisAlaGluArgAlaIlePro 110
Db 419 GAGCGCAGCCAGCGGAGAGGCCATCCCC 448

```

## RESULT 25

```

AY035778
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

AY035778
Danio rerio alpha A crystallin mRNA, complete cds.
AY035778.1 GI:18266458

```

## REFERENCE

```

AUTHORS
TITLE

```

```

Runkle,S., Hill,J., Kantorow,M., Horwitz,J. and Posner,M.
Sequence and spatial expression of zebrafish (Danio rerio)
alphaA-crystallin
Mol. Vis. 8, 45-50 (2002)

```

## MEDLINE

```

PUBMED
11925526

```

## REFERENCE

```

AUTHORS
TITLE

```

```

Submitted (16-MAY-2001) Biology, Ashland University, 401 College
Avenue, Ashland, OH 44805, USA

```

## FEATURES

```

Location/Qualifiers
1..730

```

```

/organism="Danio rerio"
/mol type="mRNA"

```

```

/db_xref="taxon:7955"

```

```

/tissue type="lens"

```

```

95..616

```

```

/note="small heat shock protein"

```

## CDS

```

/codon_start=1
/product="alpha A crystallin"
/protein_id="AAK61363.1"
/db_xref="GI:18266458"
/translation="MDIALQHPWFRRLTYTRFLDQFFGGLFDYDLFPFTTSTVSP
YRSLFRNLDSNSGVSRDEKFTVLDVKHSPEDLSVLKIDDFVEIHGHSERQDDHGYISRE
ERQDDHGYISREFHRRYRLPSNVDSATCTLSADGLTLTLCGPKTSGIDAGRGRTIP
VTREDKNSNGSSS"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 2.96e-44 Length: 730
Score: 474.00 Matches: 89
Percent Similarity: 84.92% Conservative: 18
Best Local Similarity: 70.63% Mismatches: 15
Query Match: 74.18% Indels: 4
DB: 5 Gaps: 2

```

US-10-657-740-1\_COPY\_51\_173 (1-123) x AY035778 (1-730)

```

QY 1 SerLeuPheArgThrValLeuAsp-----SerGlyIleSerGluValArgSerAsp 17
Db 239 TCACTCTTCGCAACATCTCGACTCTCCAACTCGGSGTGTCTCTGAGGTGAGTCTGAC 298
QY 18 ArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrVal 37
Db 299 AGAGAAATTTACAGTTTATCTGGATGTGAACATTTCTCTCGTATGAGTCTGAGTGC 358
QY 38 LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHis 57
Db 359 AAGGTGACAGATGACTATGTGGAGTCCAGGCAAGCATCGAAGACAGAGTATGATCAT 418
QY 58 GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSer 77
Db 419 GGTATACATCTCCGTGAGTTCCATCGTCTGCTACCGCTGCTTCCAATGTGGACAGTCT 478
QY 78 AlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGln 97
Db 479 GCATACCTGCACATGCTGCTGATGGCTCTCTCTCTTGTGGACCCCAAG---ACA 535
QY 98 ThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro 117
Db 536 AGTGGCATAGATCTGCGCGTGGAGATCGACCATCCCTGTTACCGCGGAGCAAGAGC 595
QY 118 ThrSerAlaProSerSer 123
Db 596 AACTCAGGCTCTTCTCTCC 613

```

## RESULT 26

```

BC083177
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

```

```

BC083177
Danio rerio cDNA clone MGC:92036 IMAGE:7045051, complete cds.
BC083177.1 GI:52789216
MGC.
Danio rerio (zebrafish)
Danio rerio

```

## REFERENCE

```

Actinopterygii; Neopterygii; teleostei; Ostarlophysi;
Cypriniformes; Cyprinidae; Danio.

```

## AUTHORS

```

1 (bases 1 to 1430)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.P., Jordan,H., Moore,T., Max,A.I., Wang,J., Hong,L.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
Mckernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Rulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

```

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 1430)  
Director MGC Project.  
Direct Submission  
Submitted (24-SEP-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
Tissue Procurement: Len Zon, Harvard  
CDNA Library Preparation: Open Biosystems  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Series: IRAP Plate: 178 Row: P Column: 11  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23308654.

## FEATURES

source  
1..1430  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="MGC:92036 IMAGE:7045051"  
/tissue\_type="Whole, adult zebrafish"  
/clone\_lib="NIH ZGC\_10"  
/lab\_host="DH10B"  
/note="Vector: pExpress1"  
106..627  
/codon\_start=1  
/product="Unknown (protein for MGC:92036)"  
/protein\_id="AAH83177.1"  
/db\_xref="GI:52789217"  
/translation="MDIAIQFWPFRRTGYTRLPDQFFGGLDLDLPFTTSTVSP YRSLFNLDSNGSVSEVRDEKETVLDVKHFSPELVRKVADDDYVEIHKHG ERQDDHGVISREFFHRYRLPSNVQSAITCTLSADGLLTLCGPKTSGIDAGRGRTIP VTREDDKSNSSSS"

## CDS

106..627  
/codon\_start=1  
/product="Unknown (protein for MGC:92036)"  
/protein\_id="AAH83177.1"  
/db\_xref="GI:52789217"

## ORIGIN

Alignment Scores:  
Pred. No.: 6.5e-44 Length: 1430  
Score: 474.00 Matches: 89  
Percent Similarity: 84.92% Conservative: 18  
Best Local Similarity: 70.63% Mismatches: 15  
Query Match: 74.18% Indels: 4  
DB: 5 Gaps: 2  
US-10-657-740-1\_COPY\_51\_173 (1-123) x BC083177 (1-1430)

QY 1 SerLeuPheArgThrValLeuAsp-----SerGlyIleSerGluValArgSerAsp 17  
Db 250 TCACCTCTTCCGACACATCTCGACTCTCCAACTCAGGTGCTCTGAGGTGAGGTCTGAC 309  
QY 18 ArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrVal 37  
Db 310 AGAGAAATTTACAGTTTACCTGGATGTGAACATTTCTCTCTGATGAGCTCAGTCTC 369  
QY 38 LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHis 57  
Db 370 AAGGTGACAGATGACTATGTGAGATCCAGGCGACGATGGAGAAAGACAGGATGATCAT 429  
QY 58 GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSer 77  
Db 430 GGCTACATCTCCGAGTTCATCGCCGTACCCGCTCCCTTCCAAATGTGGACCATCT 489  
QY 78 AlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGln 97  
Db 490 GCCATCCTGCACACTGTCTGCTGATGCGCTGCTCACTCTTTGTGGACCAAG---ACA 546  
QY 98 ThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluLysPro 117  
Db 547 AGTGCGATAGATGCTGGCGGTGGAGATCGCACCATCCCTGTTACCCGCGAGCAAGAC 606  
QY 118 ThrSerAlaProSerSer 123  
Db 607 AACTCGGGCTCTTCTCTCC 624  
RESULT 27  
LOCUS AY007972 676 bp mRNA linear VRT 23-OCT-2000  
DEFINITION Clarias fuscus alpha-A crystallin mRNA, complete cds.  
ACCESSION AY007972  
VERSION AY007972.1 GI:10946518  
KEYWORDS Clarias fuscus (whitespotted clarias)  
ORGANISM Clarias fuscus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes; Clariidae; Clarias.  
REFERENCE 1 (bases 1 to 676)  
AUTHORS Chlou, S.-H. and Yu, C.-M.  
TITLE Alpha crystallin of catfish eye lenses: cDNA and genomic analysis of alpha-A and alpha-B  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 676)  
AUTHORS Chlou, S.-H. and Yu, C.-M.  
TITLE Direct Submission  
JOURNAL Submitted (15-SEP-2000) Institute of Biological Chemistry, Academia Sinica, P.O. Box 23-106, Taipei, Taiwan 10098, Republic of China  
FEATURES  
source  
1..676  
/organism="Clarias fuscus"  
/mol\_type="mRNA"  
/db\_xref="taxon:33541"  
136..657  
/codon\_start=1  
/product="alpha-A crystallin"  
/protein\_id="AAG23866.1"  
/db\_xref="GI:10946519"  
/translation="MDIAIQFWPFRRTLNSPRLPDQFFGGLDLDLPFTTSTISP YRSLFRSLDSSNGSVSEVRDRPMVYLDVKHFSPELVRKVADDDYVEIHKHG ERQDDHGVISREFFHRYRLPSNVQAAITCTLSADGLLTLCGPKTGGKYGRGRTIP VTREDDKPNPAASS"

## ORIGIN

Alignment Scores:  
Pred. No.: 2.19e-43 Length: 676  
Score: 466.00 Matches: 89  
Percent Similarity: 84.13% Conservative: 17  
Best Local Similarity: 70.63% Mismatches: 16  
Query Match: 72.93% Indels: 4  
DB: 5 Gaps: 2

```

US-10-657-740-1_COPY_51_173 (1-123) x AY007972 (1-676)

QY 1 SerLeuPheArgThrValLeuAsp-----SerGlyIleSerGluValArgSerAsp 17
Db 280 TCTCTCTCCGAGATTTTGGACTCCTCAACTCTGGCATTTCCGAGGTGAGGTCTGAC 339

QY 18 ArgAspLysPheValIlePheLeuAspValIleHisGlyHisAsnGluArgGlnAspAspHis 37
Db 340 AGAGACAGATTGATGTTTACCTGGATGTGAAGCACTTCTCCAGAGGAACCTCAGATT 399

QY 38 LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHis 57
Db 400 AAGTAGCAGATGACTATCTGAGATTCTATGCAAGCAGTGGAGAAAGACAGGACGACCAT 459

QY 58 GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSer 77
Db 460 GGCTACATCTCAGCGGAGTTCCACCGCGCTACCGTCTACCTCCCAACGTCGATCAGCA 519

QY 78 AlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGln 97
Db 520 GCCATCACTTGTACGCTGTACCTGATGGCTGTCTCAGCTTCTGGGCCCCAAA---ACA 576

QY 98 ThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro 117
Db 577 GCGCGATCAAAATACGGCGTGGAGATCGCACCATTCCTGTCCACCGCGATGACAAACCC 636

QY 118 ThrSerAlaProSerSer 123
Db 637 AACCTGCTGCTTCTCTCT 654

RESULT 28
OLJ000940 436 bp mRNA linear VRT 02-MAY-1998
LOCUS Oryzias latipes, alpha-A-crystallin gene.
DEFINITION AJ000940
ACCESSION AJ000940.1 GI:3115327
VERSION alpha-A-crystallin gene.
KEYWORDS Oryzias latipes (Japanese medaka)
SOURCE Oryzias latipes
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE 1
AUTHORS Loosli,F., Koster,R.W., Carl,M., Krone,A. and Wittbrodt,J.
TITLE Six3, a medaka homologue of the Drosophila homeobox gene sine
oculus is expressed in the anterior embryonic shield and the
developing eye
JOURNAL Mech. Dev.
REFERENCE 2 (bases 1 to 436)
AUTHORS Loosli,F.
TITLE Direct Submission
JOURNAL Submitted (12-FEB-1998) Loosli F., Institute for Human Genetics,
c/o MPI of Biophys. Chem., University of Gottingen, Am Fassberg,
Gottingen, 37077, GERMANY
FEATURES
Location/Qualifiers
source
1..436
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Carolina Biological"
/db_xref="taxon:8090"
1..436
/gene="alpha-A-crystallin"
<1..>436
/gene="alpha-A-crystallin"
/protein_id="CAA04397.1"
/codon_start=1
/db_xref="GI:3115328"
/db_xref="GOA:O73919"
/db_xref="UniProt/Swiss-Prot:O73919"
translation="KLFPQFFGEGMFHDLLPFTSPITSPFYRQSLFRNFDLSSNSGI
SEVSRDRDKFTVHWDVHKHFSDELISKVVIDDFVFIQKHGERQDDHGYISREFHRRYR

LPSTWDSQAITCSLSADGLLTSLGNPAGGNRSDRSIPVCR"

ORIGIN
Alignment Scores:
Pred. No.: 4.99e-37 Length: 436
Score: 408.00 Matches: 82
Percent Similarity: 80.17% Conservative: 11
Best Local Similarity: 70.69% Mismatches: 19
Query Match: 63.95% Indels: 4
DB: Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x OLJ000940 (1-436)
QY 1 SerLeuPheArgThrValLeuAsp-----SerGlyIleSerGluValArgSerAsp 17
Db 91 TCGCTGTTTCTAATCTCTCGACTCTCAAACTCCGGCATTTCCGAGGTGAGGTCTGAC 150

QY 18 ArgAspLysPheValIlePheLeuAspValIleHisGlyLysHisAsnGluArgGlnAspAspHis 37
Db 151 AGGACAAAGTTTACCGGTTCACTGGGATGTCAAGCACTTCTCCCGATGAGCTCAGCGTG 210

QY 38 LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHis 57
Db 211 AAGGTGATCGATGACTTTGTGGAGATCCAGGCAAGCATGCGAAGAACAGACGATGACCAT 270

QY 58 GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSer 77
Db 271 GGCTACATCTCCCGGAGTTTACCGCGCTTACCGCTTCCCTCCCTCCTGTCGACCAATCG 330

QY 78 AlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGln 97
Db 331 GCCATCACTGCTCTCTGTCGCGATGGACTGTGACCTGTCCGGGCCAAACCCAGCG 390

QY 98 ThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113
Db 391 GGGGGCGCGAATGGA---CGCAGCGACGCGCAGCATCCCGTCTGCCGC 435

RESULT 29
LOCUS RATAACRYA 249 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat alpha A-crystallin mRNA, partial cds.
ACCESSION M96949
VERSION M96949.1 GI:202619
KEYWORDS alpha-A-crystallin.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 249)
AUTHORS Srinivasan,A.N., Nagineni,C.N. and Bhat,S.P.
TITLE Alpha A-crystallin is expressed in non-ocular tissues
JOURNAL J. Biol. Chem. 267 (32), 23337-23341 (1992)
MEDLINE 93054670
PUBMED 1429679
COMMENT Original source text: Rattus norvegicus (strain Sprague-Dawley)
neonate spleen cDNA to mRNA.
FEATURES
Location/Qualifiers
source
1..249
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="spleen"
/dev_stage="neonate"
1..249
/gene="alpha A-crystallin"
<1..>249
/gene="alpha A-crystallin"
/notes="this sequence is 100% homologous to previously
reported sequence of rat eye alpha A-crystallin mRNA-cDNA"
/codon_start=1
/product="alpha A-crystallin"
gene
CDS

```





```

mRNA
<120. .>419
/gene="CRYAA"
/product="alpha-A-crystallin"
120. .>419
/gene="CRYAA"
/number=3
<120. .>329
/gene="CRYAA"
/codon_start=1
/product="alpha-A-crystallin"
/db_xref="GI:2827909"
/db_xref="GI:2827909"
/translators="DNHYSIRFHHRYRLPSNVDSALSCSLADGMLTFCGPKIQT
GLDATAERAIPVSREKPTAPSS"
330. .>419
/gene="CRYAA"
3'UTR

```

```

ORIGIN
Alignment Scores:
Pred. No.: 2.14e-32 Length: 419
Score: 367.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.43% Indels: 0
DB: 9 Gaps: 0
US-10-657-740-1_COPY_51_173 (1-123) x AF026952 (1-419)

```

```

QY 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgTyrArgLeuProSerAan 73
Db 117 CAGGAGCAGCAGCGCTACATTTCCCGTGAGTTCACCGCGCTACCGCTGCCGTCAC 176
QY 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 93
Db 177 GTGGACCAAGTCGGCCCTCTCTGTCTCCCTGCTGCGATGGCATGCTGACCTTCTGTGGC 236
QY 94 ProLyIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113
Db 237 CCCAAGATCCAGACTGGCTGAGTGCACCCAGCGAGCGAGCCATCCCGCTGTGCGGG 296
QY 114 GluGluLysProThrSerAlaProSerSer 123
Db 297 GAGGAGAAGCCACCTCGGCTCCCTCGTCC 326

```

```

RESULT 32
AP001631
LOCUS
DEFINITION
Homo sapiens genomic DNA, chromosome 21, clone:KB2007G4,
AP001631
ACCESSION
AP001631.1 GI:7670585
VERSION
AP001631.1 GI:7670585
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
1 (bases 1 to 148179)
Shimizu,N., Kudoh,J. and Shibuya,K.
Homo sapiens genomic DNA, chromosome 21, clone:KB2007G4,
MX1-D21S171 region
Published Only in DataBase (2000)
2 (bases 1 to 148179)
Shimizu,N., Kudoh,J. and Shibuya,K.
Direct Submission
Submitted (04-APR-2000) Nobuyoshi Shimizu, Keio University, School
of Medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo
160-8582, Japan (E-mail:nshimizu@db.med.keio.ac.jp,
Tel:81-3-3351-2370, Fax:81-3-3351-2370)
FEATURES
Location/Qualifiers
1..148179
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

/chromosome="21"
/clone="KB2007G4"
/cell_line="FLEB14-14"
/cell_type="pre-pro-B cell"
/clone_lib="Keio BAC library"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 2.05e-29 Length: 148179
Score: 367.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.43% Indels: 0
DB: 9 Gaps: 0
US-10-657-740-1_COPY_51_173 (1-123) x AP001631 (1-148179)
QY 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgTyrArgLeuProSerAan 73
Db 132213 CAGGAGCAGCAGCGCTACATTTCCCGTGAGTTCACCGCGCTACCGCTGCCGTCAC 132272
QY 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 93
Db 132273 GTGGACCAAGTCGGCCCTCTCTGTCTCCCTGCTGCGATGGCATGCTGACCTTCTGTGGC 132332
QY 94 ProLyIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113
Db 132333 CCCAAGATCCAGACTGGCTGAGTGCACCCAGCGAGCGAGCCATCCCGCTGTGCGGG 132392
QY 114 GluGluLysProThrSerAlaProSerSer 123
Db 132393 GAGGAGAAGCCACCTCGGCTCCCTCGTCC 132422

```

## RESULT 33

```

BS000233
LOCUS
DEFINITION
Pan troglodytes chromosome 22 clone:CH251-445J22, map 22, complete
sequences.
BS000233 BA000046
ACCESSION
BS000233.1 GI:37537500
VERSION
BS000233.1 GI:37537500
KEYWORDS
HTG.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

```

## REFERENCE

```

1
The International Chimpanzee Chromosome 22 Consortium.
DNA sequence and comparative analysis of chimpanzee chromosome 22
Nature 429, 382-388 (2004)
2 (bases 1 to 163962)
Tsai,S., Liu,T., Wu,K., Liao,T. and Hsiao,K.
Direct Submission
Submitted (16-MAY-2003) Shih-Feng Tsai, National Health Research
Institutes (NHRI), Division of Molecular and Genomic Medicine; 128,
Yen-Chiu-Yuan Road, Sec 2, Taipei 115, Taiwan
(E-mail:petsai@nhri.org.tw, URL:http://www.nhri.org.tw/,
Tel:886-2-28267319, Fax:886-2-28200552)

```

## COMMENT

```

*Chinese National Human Genome Center at Shanghai, Shanghai, China;
*CBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KRIIB Genome Research
Center, Daejeon, Korea;
*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.
----- Genome Center
Center: National Yang Ming University Genome Research Center
code: YMGCC
Web site: http://genome.ym.edu.tw/
Contact: sequence@ym.edu.tw
----- Project Information

```

Center project name: The Chimpanzee Chromosome 22 Sequencing Project  
Center clone name: HX

#### ----- Summary Statistics

Sequencing vector: pUC18; 100% of reads  
Chemistry: Dye-terminator Big Dye and ET; 100% of reads Assembly  
Program: Phrap; version 0.990319  
Consensus quality: 162,928 bases at least Q40  
Consensus quality: 163,029 bases at least Q30  
Consensus quality: 163,039 bases at least Q20

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at one subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

#### ----- Source information:

The CHORI-251 chimpanzee BAC library was prepared from DNA isolated from the blood of a single male chimpanzee using published protocols (Osoegawa, K. et al. Genomics 52:1-8). The DNA from the chimpanzee ('Clint') was obtained from the Yerkes Primate Center in Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu, Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library characteristics are described at <http://www.chori.org/bacpac/chimpanzee251.htm>. The clone may be obtained from Pieter J. de Jong and coworkers (<http://www.chori.org/bacpac>).  
VECTOR: pIARBAC2.1

#### ----- Sequence Quality Assessment:

This entry has been annotated with sequence estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Neighboring clones: PTB-071B01(left) and PTB-190I13(right).

#### FEATURES

Location/Qualifiers

1..163962  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/chromosome="22"  
/clone\_lib="CH251-445322"  
/note="low quality region"  
76772..76776  
/note="gap containing unresolved di-nucleotide repeats, (CA)n"  
138531..138532  
/note="low quality region"  
138537  
/note="low quality region"  
138539..138540  
/note="low quality region"

#### ORIGIN

Alignment Scores:

Pred. No.:	2,31e-29	Length:	163962
Score:	267.00	Matches:	70
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	57.43%	Indels:	0
DB:	9	Gaps:	0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BS000233 (1-163962)

Qy 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgTyrArgLeuProSerAsn 73

Db 135934 CAGGATGACCGGCTACATTTCCTGAGTTCACCGCGCTACCGCTCCGCGCCAAAC 135993  
Qy 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyWetLeuThrPheCysGly 93  
Db 135994 GTGGACCACTGGCCCTCTCTTGCTCTGCTGCTGCGCGATGGCATGCTGACCTTCTGTGGC 136053  
Qy 94 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113  
Db 136054 CCCAGATCCAGACTGGCGCTGGATGCCACCCAGCGAGCGAGCCATCCCGCTGTGCGCG 136113  
Qy 114 GluGluLysProThrSerAlaProSerSer 123  
Db 136114 GAGGAGAAGCCACCCTCGCTCCCTCGTCC 136143

#### RESULT 34

AP001748  
LOCUS AP001748 340000 bp DNA linear PRI 21-MAY-2003  
DEFINITION Homo sapiens genomic DNA, chromosome 21q, section 92/105.  
ACCESSION AP001748 AL163293 BA000005  
VERSION AP001748.1 GI:7768745  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Hattori, M., Fujiyama, A., Taylor, T.D., Watanabe, H., Yada, T., Park, H.S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D.K., Soeda, E., Ohki, M., Takagi, T., Sakaki, Y., Taudien, S., Blechschmidt, K., Polley, A., Menzel, U., Delabar, J., Kumpf, K., Lehmann, R., Patterson, D., Reichwald, K., Rump, A., Schillhabel, M., Schudy, A., Zimmermann, W., Rosenthal, A., Sasaki, K., Nagamine, K., Kawasaki, S., Asakawa, S., Shintani, A., Sasaki, K., Kudoh, J., Shibuya, K., Mitsuyma, S., Antonarakis, S.E., Minoshima, S., Shimizu, N., Nordstok, G., Hornischer, K., Barandt, P., Scharfe, M., Schoen, O., Desario, A., Reichelt, J., Kauer, G., Bloeker, H., Ramser, J., Beck, A., Klages, S., Hennig, S., Rieselmann, L., Dagand, E., Wehrmeyer, S., Borzym, K., Gardiner, K., Nizetic, D., Francis, F., Lehrach, H., Reinhardt, R. and Yaspo, M.L.  
The DNA sequence of human chromosome 21  
Nature 405 (6784), 311-319 (2000)

#### TITLE

JOURNAL Nature 405 (6784), 311-319 (2000)  
MEDLINE 20289799  
PUBMED 10830953

#### REFERENCE

#### AUTHORS

2 (bases 1 to 340000)  
Hattori, M., Fujiyama, A., Taylor, T.D., Watanabe, H., Yada, T., Park, H.S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D.K., Soeda, E., Ohki, M., Takagi, T., Sakaki, Y., Taudien, S., Blechschmidt, K., Polley, A., Menzel, U., Delabar, J., Kumpf, K., Lehmann, R., Patterson, D., Reichwald, K., Rump, A., Schillhabel, M., Schudy, A., Zimmermann, W., Rosenthal, A., Sasaki, K., Kudoh, J., Shibuya, K., Mitsuyma, S., Asakawa, S., Shintani, A., Sasaki, T., Nagamine, K., Mitsuyma, S., Antonarakis, S.E., Minoshima, S., Shimizu, N., Nordstok, G., Hornischer, K., Barandt, P., Scharfe, M., Schoen, O., Desario, A., Reichelt, J., Kauer, G., Bloeker, H., Ramser, J., Beck, A., Klages, S., Hennig, S., Rieselmann, L., Dagand, E., Wehrmeyer, S., Borzym, K., Gardiner, K., Nizetic, D., Francis, F., Lehrach, H., Reinhardt, R. and Yaspo, M.L.

#### Direct Submission

Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing Consortium: \* RIKEN Genomic Sciences Center, Human Genome Research Group \* Institute of Molecular Biotechnology, Genome Analysis \* Keio University School of Medicine, Dept. of Molecular Biology \* GBF, Dept. of Genome Analysis \* Max-Planck Institute for Molecular Genetics (addresses see below)

#### COMMENT

On May 30, 2000 this sequence version replaced gi:7717405.  
The chromosome 21 mapping and sequencing consortium consisting of  
\* RIKEN Genomic Sciences Center, Human Genome Research Group, \* Sagami-hara 228-8555, Japan.  
\* e mail: hattori@gsc.riken.go.jp  
\* URL: <http://hgp.gsc.riken.go.jp/>  
and  
\* Institute of Molecular Biotechnology, Genome Analysis, \*

```

Beutenbergstrasse 11, D-07745 Jena, Germany,
* e.mail: gscj-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/
and
* Keio University School of Medicine, Molecular Biology, * Tokyo
160-8582, Japan,
* e.mail: nshimiz@dbm-med.keio.ac.jp
* URL: http://www.dmb.med.keio.ac.jp/
and
* GBF, Dept. of Genome Analysis,
* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
info.genome@gbf.de
* URL: http://genome.gbf.de/
and
* Max-Planck Institute for Molecular Genetics,
* Innestrasse 73, D-14195 Berlin, Germany,
* e.mail: info-chr21@molgen.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/
AL163293: Submitted (10-Apr-2000).
FEATURES
    source
        1..340000
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="21"
        <1..13998
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q22.3"
            /clone="KB51A8, 5' partial"
            /clone_lib="Keio BAC library"
            /note="Accession No. AP001628"
        5862..95299
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q22.3"
            /clone="KB1405B7"
            /clone_lib="Keio BAC library"
            /note="Accession No. AP001629"
        94424..246269
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q22.3"
            /clone="KB1151C12"
            /clone_lib="Keio BAC library"
            /note="Accession No. AP001630"
        195713..>340000
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q22.3"
            /clone="KB2007G4, 3' partial"
            /clone_lib="Keio BAC library"
            /note="Accession No. AP001631"
        1007..1195
            /note="L1M4"
            /rpt_family="LINE/L1"
            /rpt_type=DISPERSED
    repeat_region
        complement(1239..1392)
            /note="MLTIF1"
            /rpt_family="LTR/MaLR"
            /rpt_type=DISPERSED
        1606..1708
            /note="L1ME3"
            /rpt_family="LINE/L1"
            /rpt_type=DISPERSED
    repeat_region
        repeat_region
            /note="AluJo"
            /rpt_family="SINE/Alu"
            /rpt_type=DISPERSED
        2005..2209
            /note="L1ME3"
            /rpt_family="LINE/L1"
            /rpt_type=DISPERSED
        2425..2575
            /note="AluJb"
            /rpt_family="SINE/Alu"
            /rpt_type=DISPERSED
        2602..2906
            /note="AluSx"
            /rpt_family="SINE/Alu"
            /rpt_type=DISPERSED
        complement(3444..3692)
            /note="WERS8"
            /rpt_family="DNA/MER1_type"
            /rpt_type=DISPERSED
        4108..5068
            /note="CpG island"
            /rpt_type=DISPERSED
        4887..5069
            /note="(CA)n"
            /rpt_family="Simple_repeat"
            /rpt_type=TANDEM
        5105..5231
            /standard_name="WI-13399"
            /note="Accession No. G21099"
        6374..6518
            /note="MLTIF"
            /rpt_family="LTR/MaLR"
            /rpt_type=DISPERSED
        6556..6678
            /note="MLTIE"
            /rpt_family="LTR/MaLR"
            /rpt_type=DISPERSED
        8492..8788
            /note="AluSp"
            /rpt_family="SINE/Alu"
            /rpt_type=DISPERSED
        9704..9845
            /note="MLTIF"
            /rpt_family="LTR/MaLR"
            /rpt_type=DISPERSED
        10849..11001
            /note="AluSg/x"
            /rpt_family="SINE/Alu"
            /rpt_type=DISPERSED
        11600..11607
            /note="NotI site"
            /note="FAM"
            /rpt_family="SINE/Alu"
            /rpt_type=DISPERSED
        12624..12664
            /note="(CA)n"
            /rpt_family="Simple_repeat"
            /rpt_type=TANDEM
        13481..13674
            /note="HAL1"
            /rpt_family="LINE/Other"
            /rpt_type=DISPERSED
        13675..13963
            /note="AluSx"
            /rpt_family="SINE/Alu"
            /rpt_type=DISPERSED
        13964..14471
            /note="HAL1"
            /rpt_family="LINE/Other"
            /rpt_type=DISPERSED
    Alignment Scores:

```

```

Pred. No.: 5,428-29 Length: 340000
Score: 367.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.43% Indels: 0
DB: 9 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x AP001748 (1-340000)

Qy 54 GlnAspHisGlyTyrIleSerArgGluPheHisArgTyrArgLeuProSerAsn 73
Db 327926 CAGGACGACACCGGTACATTCCCGTGAGTTCCACCCCGCTACCGCTGCGTCCAAC 327985
Qy 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 93
Db 327986 GTGGACCAAGTCGGCCCTCTTGTCTCCCTGTCTGCGGATGCGATCTGACCTTCTGTGGC 328045
Qy 94 ProIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113
Db 328046 CCCAGATCCAGACGTGGCTGTGATGCCACCCAGCGGAGGAGCATCCCGTGTGCGG 328105
Qy 114 GluGluLysProThrSerAlaProSerSer 123
Db 328106 GAGGAGAGCCACCTCGGCTCCCTCGTCC 328135

RESULT 35
AJ617732
LOCUS AJ617732 507 bp mRNA linear MAM 15-JUL-2004
DEFINITION Elephas maximus partial mRNA for alphaB-crystallin (cryab gene).
ACCESSION AJ617732
VERSION AJ617732.1 GI:50344360
KEYWORDS alphaB-crystallin; cryab gene.
SOURCE Elephas maximus (Asiatic elephant)
ORGANISM Elephas maximus
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.
TITLE van Rheede, T. and de Jong, W.W.
JOURNAL The alpha-crystallins of the platypus Ornithorhynchus anatinus.
REFERENCE 2 Origin of the alternatively spliced exon alphaAins and implications
AUTHORS for mammalian phylogeny
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 507)
AUTHORS Franck, E.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-2003) Franck E., Biochemistry, Radboud University
FEATURES Location/Qualifiers
source
1..507
/organism="Elephas maximus"
/mol_type="mRNA"
/db_xref="taxon:9783"
gene 1..507
/gene="cryab"
CDS 1..507
<1..507
/gene="cryab"
/product="alphaB-crystallin"
/protein_id="CAP02108.1"
/db_xref="GI:50344361"
/translation="PWIRRRPFFPHSPSRLLFDQFGEHLLSDLPFTATSLSPFLYLR
PSFLRAPSFTGLSEMRLEKDRFSVNLVKHSPEDLKVGLGVIEVHGKHEERQD
EHGFISREPHRYKRIPADVDPLAITSSLSGDLTVNGPRKQAPGPTIPIITREPK
AVTAAPKK"
ORIGIN
Alignment Scores:
Pred. No.: 5,86e-32 Length: 507
Score: 364.00 Matches: 69
Percent Similarity: 79.31% Conservative: 23
Best Local Similarity: 59.48% Mismatches: 18
Query Match: 56.96% Indels: 6
DB: 4 Gaps: 2

```

---

```

US-10-657-740-1_COPY_51_173 (1-123) x AJ617732 (1-507)
Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 163 GACACTGGGCTCTCAGAGATCGGCTGGAGAAGACAGAGTTCTCTGTCAACCTGCGATGTG 222
Qy 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLeuHis 47
Db 223 AAGCATTTCTCTCCAGAGGATCTCAAGGTCAAGGTGCTGGGAGATGTGATGAGGTGCAT 282
Qy 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgArg 67
Db 283 GGCACATCATGAAGAACGCCAGCATGATGTTTCATCTCCCGGAGATTCCATAGGAAA 342
Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
Db 343 TACCGGATCCAGCTGATGTGACCCCTCTTGCATTTACTTCATCCCTGTCATCTCATGGG 402
Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 403 GTCTCTTACTGTGAATGGACCAAGGAAACAG-----GCCCTGGCCCTGAGCGC 450
Qy 108 AlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121
Db 451 ACCATTCCCATCACCCGGTGAAGAGAGCCTGCTGTCATCGCAGCTCCC 498

RESULT 36
AF029793
LOCUS AF029793 632 bp mRNA linear MAM 30-JUN-1999
DEFINITION Bos taurus alpha B-crystallin (CRYAB) mRNA, complete cds.
ACCESSION AF029793
VERSION AF029793.2 GI:5296002
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
TITLE Kelley, P.B., Abraham, E.C., Zhao, H.R., Shroff, N.P., Cherian, M. and
JOURNAL Direct Submission
REFERENCE 2 Submitted (10-OCT-1997) Biochemistry and Molecular Biology, Medical
AUTHORS College of Georgia, 1120 15th Street, Augusta, GA 30912, USA
2 (bases 1 to 632)
Kelley, P.B., Abraham, E.C., Zhao, H.R., Shroff, N.P., Cherian, M. and
Thomas, J.J.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1999) Biochemistry and Molecular Biology, Medical
REFERENCE 2 College of Georgia, 1120 15th Street, Augusta, GA 30912, USA
2 (bases 1 to 632)
Kelley, P.B., Abraham, E.C., Zhao, H.R., Shroff, N.P., Cherian, M. and
Thomas, J.J.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1999) Biochemistry and Molecular Biology, Medical
REMARK Sequence update by submitter
COMMENT On Jun 30, 1999 this sequence version replaced gi:2760900.
FEATURES Location/Qualifiers
source
1..632
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
gene 1..632
/gene="CRYAB"
CDS 1..528
/gene="CRYAB"
/product="alpha B-crystallin"
/protein_id="AAB95323.2"
/db_xref="GI:5296003"
/translation="MDIAIHPWIRRRPFFPHSPSRLLFDQFGEHLLSDLPFTATSL
SPFLRAPSFTGLSEMRLEKDRFSVNLVKHSPEDLKVGLGVIEVHG
KHEERQDEHGFISREPHRYKRIPADVDPLAITSSLSGDLTVNGPRKQASGPRTIP
ITREKAVTAAPKK"
ORIGIN
Alignment Scores:

```

Pred. No.:	7.58e-32	Length:	632
Score:	364.00	Matches:	58
Percent Similarity:	80.34%	Conservative:	26
Best Local Similarity:	58.12%	Mismatches:	17
Query Match:	56.96%	Indels:	6
DB:	4	Gaps:	2

US-10-657-740-1\_COPY\_51\_173 (1-123) x AF029793 (1-632)

QY	7	Leu	Asp	Ser	Gly	Ile	Ser	Glu	Val	Arg	Ser	Asp	Arg	Asp	Leu	Val	Phe	Leu	Asp	26				
Db	181	ATT	GAC	TGC	TCT	TC	GAG	ATG	GCT	CTG	GAG	AAG	GAC	AGT	TCT	CTG	TCA	ACC	TGG	240				
QY	27	Val	Lys	His	Phe	Ser	Pro	Glu	Asp	Leu	Thr	Val	Lys	Val	Gln	Asp	Asp	Phe	Val	Glu	46			
Db	241	GTT	AAG	AC	TCT	TCC	CC	CAG	GAA	CTC	AA	G	TCT	G	GG	GAG	ATG	TG	T	AGG	300			
QY	47	His	Gly	Leu	Val	His	Asn	Glu	Arg	Gln	Asp	Asp	His	Ile	Gly	Tyr	Ile	Ser	Arg	Glu	Phe	His	Arg	66
Db	301	CAT	G	CCA	AA	CAT	GA	AG	CGC	CAG	ATG	AA	CAT	G	TTT	TAT	TCC	CGG	AGT	TCC	CAC	AG	360	
QY	67	Arg	Tyr	Arg	Leu	Pro	Ser	Asn	Val	Asp	Gln	Ser	Ala	Leu	Ser	Cys	Ser	Leu	Ser	Ala	Asp	86		
Db	361	AA	TAC	CGG	ATC	CC	AG	CTG	AC	TG	AC	TG	CC	CA	TCT	ATC	CC	CT	TG	CT	CT	CG	420	
QY	87	Gly	Met	Leu	Thr	Phe	Cys	Gly	Pro	Gly	Ile	Gln	Thr	Gly	Leu	Asp	Ala	Thr	His	Ala	Glu	106		
Db	421	GGG	G	TCT	CA	CT	G	T	GA	TG	GA	CC	AA	G	AA	CA	G	-----	GC	T	CGG	CC	TG	468
QY	107	Arg	Ala	Ile	Pro	Val	Ser	Arg	Glu	Glu	Lys	Pro	-----	Thr	Ser	Ala	Pro	121						
Db	469	CGC	ACC	AT	TCC	AT	C	CCG	GA	GAG	AG	CCG	G	TG	T	C	A	CT	CG	AC	CC	CC	519	

RESULT 37  
S74229  
S74229 LOCUS  
S74229 DEFINITION  
S74229 alpha B-crystallin [rats, heart, mRNA, 671 nt].  
S74229 ACCESSION  
S74229.1 VERSION  
S74229.1 GI:241448  
S74229 671 bp mRNA linear  
S74229 ROD 07-MAY-1993

KEYWORDS	SOURCE	ORGANISM
	Rattus sp.	
	Rattus sp.	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;	
	Rattus.	
	1 (bases 1 to 671)	
REFERENCE	Atomi, Y., Yamada, S. and Nishida, T.	
AUTHORS	Early changes of alpha B-crystallin mRNA in rat skeletal muscle to	
TITLE	mechanical tension and denervation	
JOURNAL	Biochem. Biophys. Res. Commun. 181 (3), 1323-1330 (1991)	
MEDLINE	92109745	
PubMed	1764082	
REMARK	GenBank staff at the National Library of Medicine created this	
	entry [NCBI gibbsq 74229] from the original journal article.	

```

source
1. .671
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="taxon:10118"
1. .671
/gene="alpha B-crystallin"
7. .534
/gene="alpha B-crystallin"
/codon_start=1
/product="alpha B-crystallin"
/protein_id="AAB20759.1"
/db_xref="GI:241449"
/translation="MDIATHTWRPFFPHSPSRFLDQFGEHLLESDLPSTATSL
SFYLRPFSFLRASPWIDTGLSENMEKDRFSVNLVDKFSPEELKVKVLGVIEVHG
KHEERQDSHGFTISREFFHKRYIPADVDPLTIITSSLSDDGVLTVNGPRKQASGPRTIP
ITREKPAVTAAPKK"

```

## ORIGIN

CDS

Alignment Scores:	
Pred. No.:	1.78e-31
Score:	361.00
Percent Similarity:	79.4%
Best Local Similarity:	75.26%
Query Match:	56.49%
DB:	10
Length:	671
Matches:	67
Conservative:	26
Mismatches:	18
Indels:	6
Gaps:	2

US-10-657-740-1\_COPY\_51\_173 (1-123) x S74229 (1-671)

Qy	7	LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp	26
Db	187	ATTGACACTGGGCTCTCAGACATCGGTATGGAGAGACAGGTTCTCTGTGAACCTGGAC	246
Qy	27	ValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIle	46
Db	247	GTGAAGACACTTCTCTCAGAGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAGGTG	306
Qy	47	HisGlyLysHisAsnGluuArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg	66
Db	307	CACGCACAGCAGAAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGAGATTCCACAGG	366
Qy	67	ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp	86
Db	367	AAGTACCGGATCCACAGCGGACGTGGATCCTCTCACCATTACTTCTTCCCTGTCTCGGAT	426
Qy	87	GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu	106
Db	427	GGAGTCTCTCACTGTGTAATGACCACAGGAACACAG-----GCCTCTGGCCCTGAG	474
Qy	107	ArgAlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro	121
Db	475	CGCACCAATTCCCATCCCGGTGAAGAGAAGCCTGTGTCACTGCAGCCGCCCT	525

RESULT 38	RRHARTABC	LOCUS	RRHARTABC	687 bp	mRNA	linear	ROD 20-JUL-1995
					R.rattus mRNA for alpha B-crystallin (adult heart).		
					DEFINITION		
					Accession		

ACCESSION . X60352  
 VERSION X60352.1 GI:57579  
 KEYWORDS alpha-B-crystallin.  
 SOURCE Rattus rattus (black rat)  
 ORGANISM Rattus rattus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 687)  
 Bhat,S.P., Horwitz,J., Srinivasan,A. and Ding,L.  
 alphaB-crystallin exists as an independent protein in the heart and  
 in the lens  
 Eur. J. Biochem. 102, 775-781 (1991)  
 JOURNAL

REFERENCE 2 (bases 1 to 687)

```

AUTHORS      Bhat,S.P.
TITLE        Direct Submission
JOURNAL      Submitted (25-SEP-1991) S.P. Bhat,
              Institute, UCLA School of Medicine
              California 90024-7008, USA
COMMENT      A multicentric ubiquitous extracellular,
              Calmimetic related to small heat shock
              Presumed function: structural prot
              with alpha A-crystallin.
FEATURES     Location/Qualifiers
              1..687
              /organism="Rattus rattus"
              /mol_type="mRNA"
              /strain="Sprague-Dawley"
              /db_xref="taxon:10117"
              /clone="p alpha RH8"
              /cell_line="adult heart"
              /tissue_type="heart"
              /clone_lib="lambda qt11"
              /dev_stage="adult"
              ..525
              CDS

```

CDS

```

/codon_start=1
/product="alpha B-crystallin"
/protein_id="CAA42911.1"
/db_xref="GI:57580"
/db_xref="GOA:P23928"
/db_xref="UniProt/Swiss-Prot:P23928"
/translation="DIAIHHPIRRPPFPSPRLPDQPFGEHLLSDLFSTATSL
PFYLRPPSLFAPSWIDTGLSEMRMEKDRFSVNLVDVKHFSPEELKVVLGDIIEVHGK
HEERQDEHGFIISREFHRYKIPADVDPLTITSSSDGVLTVNGPRKQASGPERTIP
ITREKPAVTAAPKK"
polyA_signal
625..630
ORIGIN

Alignment Scores:
Pred. No.: 1.83e-31 Length: 687
Score: 361.00 Matches: 67
Percent Similarity: 79.49% Conservative: 26
Best Local Similarity: 57.26% Mismatches: 18
Query Match: 56.49% Indels: 6
DB: 10 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x RHARTABC (1-687)

Qy 7 LeuaspSerGlyIleSerGluValArgSerArgAspLysPheValIlePheLeuasp 26
Db 178 ATTGACCTGGCTCTCAGATGCGTATGAGAGGACAGGCTTCTGTGAACCTGGAC 237

Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspPheValGluile 46
Db 238 GTGAAGCACTTCTCCAGAGAACTCAAGTCAAGGTTCTGGGAGAGCTGATTGAGGTG 297

Qy 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrlIleSerArgGluPheHisArg 66
Db 298 CACGCAAGCAGCAGAGCGCAGCAGCAACATGCTTCTCATCTCCAGGAGTTCACACAG 357

Qy 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaasp 86
Db 358 AAGTACCGGATCCAGCGAGCTGATGCTTCTCACCATTACTTCTTCCCTGTGATCGGAT 417

Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106
Db 418 GGAGTCTCTACTGTAATGGACCAAGAAACAG-----GCCTTGGCCCTGAG 465

Qy 107 ArgAlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121
Db 466 CGCACCATTCACATCACCCTGGAAGAGAGCGCTGTGTCTGTCATGCGAGCCCT 516

RESULT 39
RLENSABC 689 bp mRNA linear ROD 31-JAN-1992
DEFINITION R.rattus mRNA for alpha B-crystallin (ocular lens tissue).
ACCESSION X60351
VERSION X60351.1 GI:57616
KEYWORDS alpha-B-crystallin.
SOURCE Rattus rattus (black rat)
ORGANISM Rattus rattus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 689)
Bhat,S.P., Horwitz,J., Srinivasan,A. and Ding,L.
alphaB-crystallin exists as an independent protein in the heart and
in the lens
Eur. J. Biochem. 102, 775-781 (1991)
2 (bases 1 to 689)
Bhat,S.P.
Direct Submission
Submitted (25-SEP-1991) S.P. Bhat, Room B-118, Jules Stein Eye
Institute, UCLA School of Medicine, 100 Stein Plaza, Los Angeles,
California 90024-7008, USA
A multimeric ubiquitous extracellular crystallin of the vertebrate
lens, related to small heat shock proteins, phosphorylated.
Presumed function: structural protein of the lens in association

FEATURES
Location/Qualifiers
1..689
/organism="Rattus rattus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10117"
/clone="p alpha RL3"
/cell_line="ocular lens tissue"
/tissue_type="lens"
/clone_lib="lambda qt11"
/dev_stage="adult"
13..540
/codon_start=1
/product="alpha B-crystallin"
/protein_id="CAA42910.1"
/db_xref="GI:57617"
/db_xref="GOA:P23928"
/db_xref="UniProt/Swiss-Prot:P23928"
/translation="MDIAIHHPIRRPPFPSPRLPDQPFGEHLLSDLFSTATSL
SPFYLRPPSLFAPSWIDTGLSEMRMEKDRFSVNLVDVKHFSPEELKVVLGDIIEVHG
KHEERQDEHGFIISREFHRYKIPADVDPLTITSSSDGVLTVNGPRKQASGPERTIP
ITREKPAVTAAPKK"
polyA_signal
555..660
ORIGIN

Alignment Scores:
Pred. No.: 1.84e-31 Length: 689
Score: 361.00 Matches: 67
Percent Similarity: 79.49% Conservative: 26
Best Local Similarity: 57.26% Mismatches: 18
Query Match: 56.49% Indels: 6
DB: 10 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x RLENSABC (1-689)

Qy 7 LeuaspSerGlyIleSerGluValArgSerArgAspLysPheValIlePheLeuasp 26
Db 193 ATTGACCTGGCTCTCAGATGCGTATGAGAGGACAGGCTTCTGTGAACCTGGAC 252

Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspPheValGluile 46
Db 253 GTGAAGCACTTCTTCCAGAGAACTCAAGTCAAGGTTCTGGGAGAGCTGATTGAGGTG 312

Qy 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrlIleSerArgGluPheHisArg 66
Db 313 CACGCAAGCAGCAGAGCGCAGCAGCAACATGCTTCTCATCTCCAGGAGTTCACACAG 372

Qy 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaasp 86
Db 373 AAGTACCGGATCCAGCGAGCTGATGCTTCTCACCATTACTTCTTCCCTGTGATCGGAT 432

Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106
Db 433 GGAGTCTCTACTGTAATGGACCAAGAAACAG-----GCCTTGGCCCTGAG 480

Qy 107 ArgAlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121
Db 481 CGCACCATTCACATCACCCTGGAAGAGAGCGCTGTGTCTGTCATGCGAGCCCT 531

RESULT 40
RLENSABC 704 bp mRNA linear ROD 06-MAY-2003
DEFINITION alpha B-crystallin [rats, heart, mRNA, 704 nt].
ACCESSION S77142
VERSION S77142.1 GI:242298
KEYWORDS Rattus sp.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 704)

```



**AUTHORS** Bhat,S.P., Horwitz,J., Srinivasan,A. and Ding,L.  
**TITLE** Alpha B-crystallin exists as an independent protein in the heart  
**JOURNAL** Eur. J. Biochem. 202 (3), 775-781 (1991)  
**MEDLINE** 92111524  
**PUBMED** 1765091  
**REMARK** GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 77142] from the original journal article.

**FEATURES** source  
 1..704  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10118"  
 <1..533  
 /codon\_start=3  
 /product="alpha B-crystallin"  
 /protein\_id="AAP31996.1"  
 /db\_xref="GI:30387800"  
 /translation="IPDIAIHPWIRRRPFPSPRLFDQFSGHLLSDFSTATSL  
 LSPFYLRPPSLRAPSWIDTGLSEMRMEKDRPSVNLVDVKHFSPELKVGLGDIIEVHG  
 GKHEERQDEHGFISREFKRYRIPADVDPLTITSSLSDDGVLTNGPRKQAGSPERTIP  
 PITREKPAVTAAPKK"

**CDS**  
 1..533  
 /product="alpha B-crystallin"  
 /protein\_id="AAP31996.1"  
 /db\_xref="GI:30387800"

## ORIGIN

**Alignment Scores:**  
 Pred. No.: 1.88e-31 Length: 704  
 Score: 361.00 Matches: 67  
 Percent Similarity: 79.49% Conservative: 26  
 Best Local Similarity: 57.26% Mismatches: 18  
 Query Match: 56.49% Indels: 6  
 DB: 10 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x S77142 (1-704)

Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 186 ATTGACATCGGGCTCTCAGAGATCGGTATGGAGAGGACAGGTTCTGTGAACCTGGAC 245  
 Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspPheValGluIle 46  
 Db 246 GTGAAGCACTTCTCCAGAGAACTCAAGTCAAGTCTCTGGAGAGCGTATGAGGTG 305  
 Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrlleSerArgGluPheHisArg 66  
 Db 306 CACGGCAAGCACGAAGAGCGCCAGACCAATCGCTTCATCTCCAGGAGTTCACAGG 365  
 Qy 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 366 AAGTACCGGATCCAGCGGACGTGGATCTCTCACCATTACTTCTTCCTGTCTATCGGAT 425  
 Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 426 GGAGTCTCTCACTGTGAATGACCAAGGAACAG-----GCCCTCGGCCCTGAG 473  
 Qy 107 ArgAlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
 Db 474 CGCACATTCCTCATCCCGTGAAGAGAGCGCTGTCTCACTGCAGGCCCT 524

**RESULT 41**  
 S77138  
**LOCUS** alpha B-crystallin [rats, lens, mRNA, 706 nt]. linear ROD 06-MAY-2003  
**DEFINITION** alpha B-crystallin [rats, lens, mRNA, 706 nt].  
**ACCESSION** S77138  
**VERSION** S77138.1 GI:242297  
**KEYWORDS**  
**SOURCE** Rattus sp.  
**ORGANISM** Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

**REFERENCE** 1 (bases 1 to 706)  
**AUTHORS** Bhat,S.P., Horwitz,J., Srinivasan,A. and Ding,L.  
**TITLE** Alpha B-crystallin exists as an independent protein in the heart

**JOURNAL** Eur. J. Biochem. 202 (3), 775-781 (1991)  
**MEDLINE** 92111524  
**PUBMED** 1765091  
**REMARK** GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 77138] from the original journal article.

**FEATURES** source  
 1..706  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10118"  
 21..548  
 /codon\_start=1  
 /product="alpha B-crystallin"  
 /protein\_id="AAP31995.1"  
 /db\_xref="GI:30387799"  
 /translation="MDIAIHPWIRRRPFPSPRLFDQFSGHLLSDFSTATSL  
 SPFYLRPPSLRAPSWIDTGLSEMRMEKDRPSVNLVDVKHFSPELKVGLGDIIEVHG  
 KHEERQDEHGFISREFKRYRIPADVDPLTITSSLSDDGVLTNGPRKQAGSPERTIP  
 ITREKPAVTAAPKK"

**CDS**  
 21..548  
 /product="alpha B-crystallin"  
 /protein\_id="AAP31995.1"  
 /db\_xref="GI:30387799"

## ORIGIN

**Alignment Scores:**  
 Pred. No.: 1.89e-31 Length: 706  
 Score: 361.00 Matches: 67  
 Percent Similarity: 79.49% Conservative: 26  
 Best Local Similarity: 57.26% Mismatches: 18  
 Query Match: 56.49% Indels: 6  
 DB: 10 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x S77138 (1-706)

Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 201 ATTGACATCGGGCTCTCAGAGATCGGTATGGAGAGGACAGGTTCTGTGAACCTGGAC 260  
 Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspPheValGluIle 46  
 Db 261 GTGAAGCACTTCTCCAGAGAACTCAAGTCAAGTCTCTGGAGAGCGTATGAGGTG 320  
 Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrlleSerArgGluPheHisArg 66  
 Db 321 CACGGCAAGCACGAAGAGCGCCAGACCAATCGCTTCATCTCCAGGAGTTCACAGG 380  
 Qy 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 381 AAGTACCGGATCCAGCGGACGTGGATCTCTCACCATTACTTCTTCCTGTCTATCGGAT 440  
 Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 441 GGAGTCTCTCACTGTGAATGACCAAGGAACAG-----GCCCTCGGCCCTGAG 488  
 Qy 107 ArgAlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
 Db 489 CGCACATTCCTCATCCCGTGAAGAGAGCGCTGTCTCACTGCAGGCCCT 539

**RESULT 42**  
 AX401743  
**LOCUS** Sequence 1419 from Patent WO0210453. linear PAT 06-JUN-2002  
**DEFINITION** Sequence 1419 from Patent WO0210453.  
**ACCESSION** AX401743  
**VERSION** AX401743.1 GI:21337923  
**KEYWORDS**  
**SOURCE** Rattus norvegicus (Norway rat)  
**ORGANISM** Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

**REFERENCE** 1  
**AUTHORS** Mendrick,D., Porter,M.W., Johnson,K.R., Castle,A.L. and  
 Elashoff,M.R.  
**TITLE** Molecular toxicology modeling  
**JOURNAL** Patent: WO 0210453-A 1419 07-FEB-2002;





```

QY 79 LeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThr 98
   ::::: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 397 ATTACTTCATCCCTGCTCATCTGATGGGGTCTCTCACTGTGAATGACCAAGAAACAG--- 453

QY 99 GlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro 117
   ::::: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 454 -----GCCTCTGGCCCTGAGGCTACCAITCCATCCACACGTCGAGAAAGGCT 501

RESULT 46
SJU293658          528 bp  mRNA  linear  ROD 20-MAR-2001
LOCUS      Spalax judaei mRNA for alpha-B-crystallin.
DEFINITION  AJ293658
ACCESSION  AJ293658
VERSION    AJ293658.1 GI:9716998
KEYWORDS   alpha-B-crystallin.
SOURCE     Spalax judaei
ORGANISM   Spalax judaei
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Spalacinae;
            Spalax.
REFERENCE   1
AUTHORS    Avioli, A., Joel, A. and Nevo, E.
TITLE      The lens protein alpha-B-crystallin of the blind subterranean
            mole-rat: high homology with sighted mammals
JOURNAL    Gene 264 (1), 45-49 (2001)
MEDLINE    21142392
PUBMED     11245977
REFERENCE   2 (bases 1 to 528)
AUTHORS    Avioli, A.
DIRECT SUBMISSION
SUBMITTED  (02-AUG-2000) Avioli A., Institute of Evolution, Haifa
UNIVERSITY, Mt. Carmel, Haifa, 31905, ISRAEL
LOCATION/Qualifiers
FEATURES   1..528
            /organism="Spalax judaei"
            /mol_type="mRNA"
            /isolate="Anza population"
            /db_xref="taxon:134510"
            /clone="g60aBcrystal"
            /sex="female"
            /tissue_type="heart and lens"
            /dev_stage="adult"
            /country="Israel:Samaria Mt."
            /note="Spalax Judaei belongs to the Spalax Ehrenbergi
            superspecies, the name was not validly published at the
            time of sequence submission"
            1..528
            /gene="alpha-B-crystallin"
            1..528
            /gene="alpha-B-crystallin"
            /function="a major protein in the eye lens"
            /codon_start=1
            /evidence=experimental
            /product="alpha-B-crystallin"
            /protein_id="CAC01692.1"
            /db_xref="GI:9716998"
            /db_xref="UniProt/TREMBL:Q9BPF3"
            /translation="MDIALHHPWRPPFPFPHSPSRLEDFQPFGEHLLSDFSTSTSL
            SPFYLRPSFFRAPSWDITGLSEMRKRLSLNDVVKHFSPEELKVKVLGVDVIEVHG
            KHEERODSHGFTISRFHKYRIPADVDPLTITSSLSDDGLTVNGPRKQASGPENTIP
            ITREKPAVTAAPKK"
            source
gene
CDS
ORIGIN
Alignment Scores:
Pred. No.: 3..82e-31 Length: 528
Score: 357.00 Matches: 69
Percent Similarity: 76.80% Conservative: 27
Best Local Similarity: 55.20% Mismatches: 21
Query Match: 55.87% Indels: 8
DB: 10 Gaps: 3
US-10-657-740-1_COPY_51_173 (1-123) x SJU293658 (1-528)

```

```

QY 1 SerLeuPheArg-----ThrValLeuAspSerGlyIleSerGluValArgSerAspArg 18
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 157 TCCTTCTTTTCGGGCACCCAGCTGGATTGACACTGCACTCTCAGAGATGCGTATGGAGAAG 216

QY 19 AspLysPheValIlePheIleuAspValIleHisPheSerProGluAspLeuThrValLys 38
   ::::: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 217 GACAGGCTCTCTGCTCAACCTGGATGTGAAGCACTTCTCCCCAGAGAACTCAAGGTCAAG 276

QY 39 ValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspHisGly 58
   ::::: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 277 GTGTGGGAGATGTGATTGAGTGCACGCGCAAGCATGAAGACGCGGATGAGCATGCT 336

QY 59 TyrIleSerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAla 78
   ::::: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 337 TTCATCTCCAGAGAGTTCCACCGGAAGTACCGGATCCAGCTGTGTGGACCTCTCACT 396

QY 79 LeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThr 98
   ::::: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 397 ATTACTTCATCCCTGCTCATCTGATGGGGTCTCTCACTGTGAATGGAACCAAG--- 453

QY 99 GlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro--- 117
   ::::: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 454 -----GCCTCTGGCCCTGAGGCTACCAITCCATCCACACGTCGAGAAAGGCTGCT 504

QY 118 ---ThrSerAlaPro 121
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 505 GTCACCGCGGCCCT 519

RESULT 47
RCABCRYST
LOCUS      R.Catesbeiana mRNA for alpha-B-crystallin.
DEFINITION  X87114
ACCESSION  X87114
VERSION    X87114.1 GI:809568
KEYWORDS   alpha-B-crystallin.
SOURCE     Rana catesbeiana (bullfrog)
ORGANISM   Rana catesbeiana
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana;
            Aquarana.
REFERENCE   1
AUTHORS    Lu, S.F., Pan, F.M. and Chiou, S.H.
TITLE      Sequence analysis of frog alpha B-crystallin cDNA: sequence
            homology and evolutionary comparison of alpha A, alpha B and heat
            shock proteins
JOURNAL    Biochem. Biophys. Res. Commun. 216 (3), 881-891 (1995)
MEDLINE    96074566
PUBMED     7488208
REFERENCE   2 (bases 1 to 522)
AUTHORS    Lu, S.F.
DIRECT SUBMISSION
SUBMITTED  (09-MAY-1995) S.F. Lu, Institute of Biological Chemistry,
UNIVERSITY, PO Box 23-106, Taipei 10098, TAIWAN
LOCATION/Qualifiers
FEATURES   1..522
            /organism="Rana catesbeiana"
            /mol_type="mRNA"
            /db_xref="taxon:8400"
            /tissue_type="lens"
            1..522
            /gene="alpha-B-crystallin"
            1..522
            /gene="alpha-B-crystallin"
            /codon_start=1
            /product="alpha-B-crystallin"
            /protein_id="CAA60594.1"
            /db_xref="GI:809569"
            /db_xref="GOA:Q91312"
            /db_xref="UniProt/Swiss-Prot:Q91312"
            /translation="MDITQHPFRFRQYSPFGPNKMFQCFGEHQIADLPFSPSVLS
            PFYKYPFLRLPWSWIESGLSEMRLEKDKFSLNDVVKHFSPEELKVKVSGDFTIEHGKH

```



```

source
1. 528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GH00548X1.0"
/clone_lib="BD Creator (TM) CDS Library derived from MGC
collection"
/lab_host="DHSalpha T1 resistant"
/note="Vector: pDNR-Dual"
1. 528
CDS
/product="crystallin, alpha B"
/protein_id="AAP35416.1"
/db_xref="GI:30582379"
/translating="MDIAIHHPWIRPPFPHPSPRLPDPQFFGHEHLLSDLPFTSTSL
SPFYLRPSFLRSPFDTGLSEMRLEKDFSVNLDVKHFSPEELKVKVLGDVIEVHG
KHEERQDEHGFISREFHRYIPADVDPLTITSSLSGDLTVNGPRKQVSGPERTIP
ITREKPAVTAAPKK"

ORIGIN
Alignment Scores:
Pred. No.: 8,37e-31 Length: 528
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservat: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 9 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x BT006770 (1-528)

Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 184 GACACTGGACTCTCAGAGATCGCGTGGAAAGGACAGGTTCTCTGTCACCTGGATGTG 243
Qy 28 LysHisPheSerProGluAspLeuThrValIleValGlnAspAspPheValGluLeuHis 47
Db 244 AAGCACTTCTCCAGAGAACTCAAGTTAAGGTGTTGGAGATGTTGATGAGTGCAT 303
Qy 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrlleSerArgGluPheHisArgArg 67
Db 304 GGAACAACATGAAGAGCCGAGATGATGTTTCACTCCAGGAGTTTCCACAGGAAA 363
Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
Db 364 TACCGGATCCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 423
Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 424 GTCTCTCACTGTAATGACCAAGAAACAG-----GTCTCTGGCCCTGAGCGC 471

RESULT 50
BT007909
LOCUS
DEFINITION
BT007909 528 bp mRNA linear SYN 13-MAY-2003
Synthetic construct Homo sapiens crystallin, alpha B mRNA, partial
cde.
ACCESSION
BT007909.1 GI:30584656
KEYWORDS
FLI CDNA.
SOURCE
synthetic construct
synthetic construct
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 528)
AUTHORS
Kalinina, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,
Koundinya, M., Raphael, J., Moreira, D., Kelley, T., LaBaer, J., Lin, Y.,
Phelan, M. and Farmer, A.
TITLE
Cloning of human full-length CDSs in BD Creator (TM) System Donor
vector
JOURNAL
unpublished
REFERENCE
2 (bases 1 to 528)

```

```

AUTHORS
Kalinina, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,
Koundinya, M., Raphael, J., Moreira, D., Kelley, T., LaBaer, J., Lin, Y.,
Phelan, M. and Farmer, A.
TITLE
Direct Submission
JOURNAL
Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USA
COMMENT
This CDS clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the
Harvard Institute of Proteomics. Each CDS has been cloned in two
forms: with and without stop-codon (to allow fusion with C-terminal
tag). The CDS has been directionally cloned using BD In-Fusion(TM)
cloning system between the SalI and HindIII sites of the pDNR-DUAL
vector. Additional sequences in the clone: 'ACC' after SalI site
and before 'ARG' to provide kozak consensus sequence; 'GG' after
last codon and before HindIII site to maintain reading frame.
Clone distribution: http://bioinfo.clontech.com/orfclones.
FEATURES
Location/Qualifiers
1..528
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="GH00548L1.0"
/clone_lib="BD Creator (TM) CDS Library derived from MGC
collection"
/lab_host="DHSalpha T1 resistant"
/note="Vector: pDNR-Dual"
1..528
CDS
/note="Mutations: 527:Stop->Leu"
/codon_start=1
/translating="Homo sapiens crystallin, alpha B"
/product="Homo sapiens crystallin, alpha B"
/protein_id="AAP36581.1"
/db_xref="GI:30584657"
/translating="MDIAIHHPWIRPPFPHPSPRLPDPQFFGHEHLLSDLPFTSTSL
SPFYLRPSFLRSPFDTGLSEMRLEKDFSVNLDVKHFSPEELKVKVLGDVIEVHG
KHEERQDEHGFISREFHRYIPADVDPLTITSSLSGDLTVNGPRKQVSGPERTIP
ITREKPAVTAAPKKL"

ORIGIN
Alignment Scores:
Pred. No.: 8,37e-31 Length: 528
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservat: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 12 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x BT007909 (1-528)

Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 184 GACACTGGACTCTCAGAGATCGCGTGGAAAGGACAGGTTCTCTGTCACCTGGATGTG 243
Qy 28 LysHisPheSerProGluAspLeuThrValIleValGlnAspAspPheValGluLeuHis 47
Db 244 AAGCACTTCTCCAGAGAACTCAAGTTAAGGTGTTGGAGATGTTGATGAGTGCAT 303
Qy 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrlleSerArgGluPheHisArgArg 67
Db 304 GGAACAACATGAAGAGCCGAGATGATGTTTCACTCCAGGAGTTTCCACAGGAAA 363
Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
Db 364 TACCGGATCCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 423
Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 424 GTCTCTCACTGTAATGACCAAGAAACAG-----GTCTCTGGCCCTGAGCGC 471

TITLE
Cloning of human full-length CDSs in BD Creator (TM) System Donor
vector
JOURNAL
unpublished
REFERENCE
2 (bases 1 to 528)

```

Tue May 31 10:29:17 2005

us-10-657-740-1\_copy\_51\_173.rge

Page 32

Search completed: May 30, 2005, 07:16:48  
Job time : 3109.55 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 29, 2005, 22:54:38 ; Search time 374.402 Seconds

(without alignments)  
1944.776 Million cell updates/sec

Title: US-10-657-740-1\_COPY\_51\_173

Perfect score: 639

Sequence: 1 SLFRVLDGISEVSRDRDK.....HAERAIPVSRREKPTAPSS 123

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10657740/runat\_27052005\_165252\_3305/app\_query.fasta\_1.590  
-DB=N Geneseq 16Dec04 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=150 -DOCALLIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=50  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10657740 @CGN 1 1 796 @runat\_27052005\_165252\_3305 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DISPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 16Dec04: \*  
1: geneseqm1980s: \*  
2: geneseqm1990s: \*  
3: geneseqm2000s: \*  
4: geneseqm2001as: \*  
5: geneseqm2001bs: \*  
6: geneseqm2002as: \*  
7: geneseqm2002bs: \*  
8: geneseqm2003as: \*  
9: geneseqm2003bs: \*  
10: geneseqm2003cs: \*  
11: geneseqm2003ds: \*  
12: geneseqm2004as: \*  
13: geneseqm2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	639	100.0	372	13	ADQ78288 Human N-t
2	599	93.7	1056	10	AD52521 Primary r
3	599	93.7	1056	10	ADF30549 Rat angio
4	577.5	90.4	1271	10	ADF30547 Rat angio
5	567	88.7	531	8	ABX12062 Human alp

6	367	57.4	211	12	ACH93386	Ach93386 Human gen
7	367	57.4	573	12	ACH79686	ACH79686 Human gen
8	361	56.5	528	10	ADB52642	ADB52642 Primary r
9	361	56.5	1247	10	ABT41872	ABT41872 Toxicity
10	361	56.5	1247	6	ABK63512	ABK63512 Rat seque
11	361	56.5	1247	12	ADP72683	ADP72683 Renal tox
12	358	56.0	548	6	ABZ35138	ABZ35138 Human gen
13	354	55.4	528	13	ADR46388	ADR46388 Human ten
14	354	55.4	537	10	ADE75374	ADE75374 Human mit
15	354	55.4	537	10	ADE75375	ADE75375 Human wil
16	354	55.4	691	2	AA339668	AA339668 Renal can
17	354	55.4	691	2	ABL65204	ABL65204 Lung canc
18	354	55.4	691	6	ABL62427	ABL62427 Colon ade
19	354	55.4	691	6	ABT10966	ABT10966 Human bre
20	354	55.4	691	6	ABK64671	ABK64671 Human ben
21	354	55.4	691	12	ADF09686	ADF09686 Human cry
22	354	55.4	691	10	ADN05800	ADN05800 Antipsori
23	354	55.4	691	13	ADR24825	ADR24825 Breast ca
24	354	55.4	856	3	AAC03893	AAC03893 Human sec
25	354	55.4	893	3	AAC10867	AAC10867 Human sec
26	354	55.4	913	3	AAC10863	AAC10863 Human sec
27	354	55.4	913	3	AAC10862	AAC10862 Human sec
28	354	55.4	927	3	AAC10861	AAC10861 Human sec
29	354	55.4	942	3	AAC03892	AAC03892 Human sec
30	354	55.4	1036	2	AAC23574	AAC23574 Human bre
31	332.5	52.0	380	8	ABX38978	ABX38978 Bovine ES
32	330	51.6	695	3	AAC03894	AAC03894 Human sec
33	283	44.3	449	8	ABX39112	ABX39112 Bovine ES
34	263	41.2	349	6	ABK87386	ABK87386 Target mo
35	261.5	40.9	615	12	ADN31647	ADN31647 Human hep
36	261.5	40.9	764	12	ADM94741	ADM94741 Human hea
37	261.5	40.9	789	11	ADI31974	ADI31974 Human cdn
38	261.5	40.9	847	10	ADG10693	ADG10693 Human STA
39	261.5	40.9	847	13	ADQ86169	ADQ86169 Human tum
40	261.5	40.9	865	6	ABO60780	ABO60780 Human HSB
41	261.5	40.9	865	13	ADR24765	ADR24765 Breast ca
42	261.5	40.9	865	13	ACN38656	ACN38656 Tumour-as
43	261.5	40.9	865	13	ADP23292	ADP23292 PRO polyo
44	261.5	40.9	1231	6	ABN97370	ABN97370 Gene #386
45	261.5	40.9	1231	6	ABN97370	ABN97370 Human hea
46	261.5	40.9	1380	3	AAA93441	AAA93441 GFP-HSP27
47	261.5	40.9	1380	6	AB871560	AB871560 DNA encod
48	261	40.8	604	6	ABK63074	ABK63074 Rat seque
49	261	40.8	604	12	ADP72259	ADP72259 Renal tox
50	260	40.7	599	2	ABV94670	ABV94670 Human pan
51	260	40.7	618	2	AAQ48718	AAQ48718 Encodes p
52	260	40.7	787	6	ABK63521	ABK63521 Rat seque
53	260	40.7	787	10	ADB58686	ADB58686 Toxicity
54	260	40.7	787	10	ABT42261	ABT42261 Toxicity
55	260	40.7	787	12	ADP72725	ADP72725 Renal tox
56	260	40.7	787	13	ADT66606	ADT66606 Rat heat
57	260	40.7	1379	3	AAC66073	AAC66073 E. coli e
58	260	40.7	1379	4	AAF77805	AAF77805 Fragment
59	260	40.7	1379	4	AAF61192	AAF61192 Expressio
60	260	40.7	1379	5	AAD02211	AAD02211 Partial p
61	259	40.5	396	8	ABX47457	ABX47457 Bovine ES
62	258	40.4	719	4	ABL07843	ABL07843 Drosophi
63	256	40.1	560	10	ADK12087	ADK12087 Breast ca
64	250.5	38.2	480	10	ADH76210	ADH76210 Human hea
65	250.5	38.2	486	10	ADH76202	ADH76202 Rat heat
66	250.5	38.2	575	6	ABQ61117	ABQ61117 Skeletal
67	250.5	38.2	1310	10	ABT41710	ABT41710 Toxicity
68	250.5	39.2	1433	5	ADM19230	ADM19230 Novel hum
69	250.5	39.2	1488	2	AAX24230	AAX24230 Human nor
70	250.5	39.2	1506	5	ADM19501	ADM19501 Novel hum
71	250.5	39.2	1771	3	AAC76681	AAC76681 Human ORP
72	250.5	39.2	1820	12	ADQ22255	ADQ22255 Human sof
73	250	39.1	450	9	ACH17737	ACH17737 Human adu
74	248	38.8	503	6	ABL99507	ABL99507 Target ca
75	248	38.8	558	10	ADK12085	ADK12085 Breast ca
76	242.5	37.9	909	2	AAQ03969	AAQ03969 Clone P25
77	242	37.9	543	12	ADL11837	ADL11837 Cat flea
78	241.5	37.8	708	5	ADL40880	ADL40880 Human ova

c	79	240.5	37.6	376	4	AAL19387	Aa119387 Human bre
	80	240.5	37.6	411	11	ACN89210	ACN89210 Breast ca
	81	240.5	37.6	722	13	ADT66608	ADT66608 Murine hé
c	82	240.5	37.6	563	11	ACN87226	ACN87226 Breast ca
	83	234.5	36.7	614	5	ADI75644	ADI75644 Human ova
	84	234.5	36.7	614	5	ADI69298	ADI69298 Human ova
	85	233.5	36.5	755	2	AAQ03970	AAQ03970 Clone P25
	86	233.5	36.5	954	5	ADL62061	ADL62061 Human ova
	87	229	35.8	1700	4	AA556378	AA556378 Human cDN
	88	226	35.4	421	9	ACH17560	ACH17560 Human adu
	89	223.5	35.0	500	9	ACH34922	ACH34922 Human end
	90	222	34.7	405	9	ACH16863	ACH16863 Human adu
c	91	216	33.8	2874	4	ABL07842	ABL07842 Drosophil
c	92	216	33.8	3119	4	ABL09418	ABL09418 Drosophil
c	93	212	33.2	49634	6	ABL68447	ABL68447 Kidney ca
	94	211.5	33.1	870	4	ABL68447	ABL68447 Kidney ca
c	95	211.5	33.1	2870	4	ABL07721	ABL07721 Drosophil
	96	207	32.4	387	12	ADQ17563	ADQ17563 Human sof
	97	202.5	31.7	937	4	ABL07315	ABL07315 Drosophil
	98	202.5	31.7	2651	4	ABL07714	ABL07714 Drosophil
	99	202.5	31.7	2937	4	ABL07314	ABL07314 Drosophil
	100	202	31.6	212	8	ABT33639	ABT33639 Anticance
c	101	198.5	31.1	396	6	ABL63575	ABL63575 Breast ca
c	102	198.5	31.1	396	6	ABL63575	ABL63575 Breast ca
c	103	197.5	30.9	549	3	AC93846	AC93846 Cat flea
	104	195	30.5	467	6	ABV94661	ABV94661 Human pan
	105	193	30.2	600	4	ABL07321	ABL07321 Drosophil
	106	193	30.2	2600	4	ABL07320	ABL07320 Drosophil
	107	192.5	30.1	482	9	ACH17221	ACH17221 Human adu
	108	192.5	30.1	550	12	ADL11752	ADL11752 Cat flea
	109	192	30.0	6806	9	ACF25379	ACF25379 Rat alpha
	110	188.5	29.5	1206	4	ABL07725	ABL07725 Drosophil
c	111	188.5	29.5	3206	4	ABL07724	ABL07724 Drosophil
	112	188	29.4	4206	6	ABZ35176	ABZ35176 Human gen
	113	188	29.4	4206	10	ADE84874	ADE84874 Farnesyl
	114	182.5	28.6	393	3	AC06428	AC06428 Human sec
	115	182.5	28.6	410	6	ABS70907	ABS70907 Deer cDNA
c	116	179	28.0	218	4	AAL09000	AAL09000 Human bre
c	117	179	28.0	565	11	ACN79216	ACN79216 Breast ca
	118	174.5	27.3	1367	4	ABL07295	ABL07295 Drosophil
	119	174.5	27.3	3367	4	ABL07294	ABL07294 Drosophil
c	120	174	27.2	510	6	ABQ14062	ABQ14062 Oligonuc
	121	174	27.2	510	6	ABQ14063	ABQ14063 Oligonuc
	122	172	26.9	542	12	ADL11704	ADL11704 Cat flea
	123	170	26.6	654	5	AA542494	AA542494 Human cDN
	124	165.5	25.9	487	13	ADS54436	ADS54436 Bacterial
	125	165.5	25.9	487	13	ADS54437	ADS54437 Bacterial
	126	163.5	25.6	1811	12	ADJ75836	ADJ75836 Marker ge
	127	163	25.5	445	9	ACH23297	ACH23297 Human adu
	128	161.5	25.3	651	4	ABL07715	ABL07715 Drosophil
c	129	161.5	25.3	2651	4	ABL07714	ABL07714 Drosophil
c	130	161.5	25.3	2937	4	ABL07314	ABL07314 Drosophil
	131	160	25.0	1495	10	ADI22459	ADI22459 Rat liver
	132	159	24.9	1627	2	AAZ21959	AAZ21959 cDNA enco
	133	159	24.9	1627	9	ACH04117	ACH04117 Human cDN
	134	159	24.9	1843	4	AA743328	AA743328 Human H11
	135	159	24.9	1843	4	AA743307	AA743307 Human H11
	136	159	24.9	1854	5	ABX71326	ABX71326 Human met
	137	159	24.9	2004	5	ADM19232	ADM19232 Novel hum
	138	159	24.9	2010	12	ADJ74941	ADJ74941 Marker ge
	139	159	24.9	2010	13	ADR25621	ADR25621 Breast ca
c	140	159	24.9	2012	4	AAK52551	AAK52551 Human pol
	141	159	24.9	2036	10	ADI22637	ADI22637 Human liv
	142	159	24.9	2048	4	AAK51567	AAK51567 Human pol
	143	158	24.7	764	6	ABX84243	ABX84243 Human cDN
	144	158	24.7	810	12	ADQ22105	ADQ22105 Human sof
	145	154	24.1	823	2	AAZ24422	AAZ24422 Human bla
	146	151.5	23.7	1176	4	ABL24579	ABL24579 Drosophil
	147	150	23.5	438	11	ADN31653	ADN31653 Nematode
	148	149.5	23.4	333	3	AAO6427	AAO6427 Human sec
	149	148	23.2	403	8	ABX38783	ABX38783 Bovine ES
	150	148	23.2	490	13	ADS54435	ADS54435 Bacterial

ALIGNMENTS

RESULT 1

ADQ78288 standard; cDNA; 372 BP.

XX, ADQ78288;

XX 04-NOV-2004 (first entry)

XX Human N-terminal truncated alpha-crystallin DNA.

XX alpha-crystallin; enhanced protein expression;

XX enhanced protein secretion; protein aggregation; heat tolerance;

XX elevated temperature; human; ds; gene.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..372

XX /partial

XX /tag= a

XX /product= "N-terminal truncated alpha-crystallin"

XX /note= "No start codon given"

XX US2004157289-A1.

XX 12-AUG-2004.

XX 08-SEP-2003; 2003US-00657740.

XX 06-SEP-2002; 2002US-0408680P.

XX (SALE/) SALERNO J C.

XX (HANN/) HANNA M.

XX (KORE/) KORETZ J F.

XX (CRON/) CRONE D.

XX (SMIT/) SMITH S M E.

XX Salerno JC, Hanna M, Koretz JF, Crone D, Smith SME;

XX WPI; 2004-580268/56.

XX P-PSDB; ADQ78289.

XX New truncated alpha-crystallin polypeptide derived from a wild-type alpha

XX -crystallin protein, useful for enhancing protein (e.g. insulin or

XX alcohol dehydrogenase) expression or secretion and for preventing protein

XX aggregation.

XX Claim 16; SEQ ID NO 2; 33pp; English.

XX The invention relates to a truncated alpha-crystallin polypeptide derived

XX from a wild-type alpha-crystallin protein, where the truncated

XX polypeptide lacks an N-terminal sequence present in the wild-type

XX protein. The composition and methods are useful for enhancing protein

XX (e.g. insulin or alcohol dehydrogenase) expression or secretion and for

XX preventing protein aggregation. These may also be used for creating a

XX thermophilic host that tolerates elevated temperatures. The present

XX sequence represents human N-terminal truncated alpha-crystallin DNA.

XX SQ Sequence 372 BP; 66 A; 136 C; 98 G; 72 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,07e-75 Length: 372

Score: 639.00 Matches: 123

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 13 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x ADQ78288 (1-372)





Qy	18	ArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrVal	37
Db	429	CGGACAAGATTGTTCATCTTCTGTGATGTGAAGCACTTCTCTCTGAGACCTCACCGTG	488
Qy	38	LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspHis	57
Db	489	AAGGTACTCGGAAGATTTCGTGAGATCCATGTCGAACAACAACAGAGAGGCAGGATGCACAT	548
Qy	58	GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSer	77
Db	549	GGCTACATTTCCGTGAATTTCCACGTCGCTACCGCTCTGCTCTCCATGTGACCAAGTCC	608
Qy	78	AlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGln	97
Db	609	GCCCTCTCCTGTCTCTGTCTGCGGATGCATGCTGCACCTTCTCTGGCCCCAAGGTCCAG	668
Qy	98	ThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro	117
Db	669	TCTGGCTTGGATGCTGGCCACAGCGAGAGGGCCATTCCCGTGTACGGGAGGAGAGGCC	728
Qy	118	ThrSerAlaProSerSer	123
Db	729	AGCTCGGACCCCTCGTCC	746

## RESULT 5

RESOL 3  
ABX12062  
ID ABX12062 standard; DNA: 531 BP.

XX  
DT 16-MAY-2003 (first entry)

DE Human alpha BNAC crystallin chimera DNA.

Human; ds; gene; chimera; alpha BNC crystallin; protein shelf life; protein aggregation; accessible hydrophobic region increase; mutant; larger size oligomer formation; intersubunit interaction increase; larger aggregate formation; larger porous oligomer formation; increased ellipticity; less solvent accessible tryptophan; increased chaperone-like activity; alpha A crystallin; alpha B crystallin.

OS	Homo sapiens.
OS	Synthetic.

Key	Location/Qualifiers
CDS	1..531
FT	/tag= a
FT	/product= "Alpha BNAC crystallin"
FT	/partial
FT	/note= "No stop codon given"
FT	1..247
FT	/tag= b
FT	/note= "Derived from DNA sequence
FT	sequence of alpha B crystallin"
FT	248..531
FT	/tag= c
FT	/note= "Derived from DNA sequence
FT	sequence of alpha A crystallin"

PN US2002177192-A1.

28-NOV-2002

26-MAR-2002: 2002US-00105427.

PR 28-MAR-2001: 2001US-0279223P.

PA (KUMA//) KUMAR L V S.  
PA (RAOC//) RAO C M.

XX  
PI  
Kumar LVS, Rao CM:XX  
TJ

DR WPI; 2003-298776/29.  
DR P-PSDB; ABG76084.

New chimera alpha BNAC nucleic acid, useful for preventing aggregation of PT proteins and also for increasing shelf life of proteins of pharmaceutical PT value.

PS Claim 1; Fig 9; 17pp; English.

The invention relates to a chimera alpha BNAC polynucleotide that encodes a chimera alpha BNAC polypeptide. The polypeptide is useful for preventing protein aggregation. The polypeptide is also useful for increasing the shelf life of proteins of pharmaceutical value. The polypeptide shows an increase in accessible hydrophobic regions, forms larger size oligomers, shows an increase in intersubunit interaction, forms larger aggregates, forms larger porous oligomers and shows increased ellipticity as compared to eye lens crystallins alpha A and alpha B. The tryptophan residues in the polypeptide are less solvent accessible as compared to those of eye lens crystallins alpha A and alpha B. The polypeptide shows extraordinarily high chaperone-like activity ranging between 3 - 6 times that of the eye lens crystallins alpha A and alpha B. The present sequence represents the chimera DNA that encodes human alpha BNAC crystallin

Sequence 531 BP; 90 A; 198 C; 130 G; 113 T; 0 U; 0 Other; XX

**Alignment Scores:**

Pred. No.:	4,416-65	Length:	531
Alignment Scores:		Matches:	106
Score:	\$67.00	Conservative:	7
Percent Similarity:	97.41%	Mismatches:	3
Best Local Similarity:	91.38%	Indels:	0
Query Match:	88.73%	Gaps:	0
DB:	8		

US-10-657-740-1 COPY 51 173 (1-123) x ABX12062 (1-531)

[illegible]

RESULT 6  
ACH93386

ACH93386  
ID ACH93386 standard: DNA: 211 bp.

XX  
ACH93386:XX  
DT 29-JUL-2004 (first entry)

XX Human genome derived single exon probe #26581.

[illegible]

Human; probe; ss; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.

XX  
NY

OS Homo sapiens.  
 XX US2003194704-A1.  
 XX 16-OCT-2003.  
 XX 03-APR-2002; 2002US-00029386.  
 XX 03-APR-2002; 2002US-00029386.  
 XX (PENN/) PENN S G.  
 XX (RANK/) RANK D R.  
 XX (HANZ/) HANZEL D K.  
 XX Penn SG, Rank DR, Hanzel DK;  
 XX WPI; 2004-119264/12.  
 XX New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for  
 PT surveying tissues.  
 XX Claim 1; SEQ ID NO 26581; 80pp; English.  
 XX The invention relates to a nucleic acid probe for measuring human gene  
 CC expression, comprising any of the 27,400 fully defined nucleotide  
 CC sequences in the specification, or their complements or fragments, and  
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
 CC fully defined in the specification. The probe is a single exon probe that  
 CC hybridises under high stringency conditions to a nucleic acid molecule  
 CC expressed in human cells or tissues. Also included are a spatially-  
 CC addressable set of single exon nucleic acid probes for measuring human  
 CC gene expression (comprising a plurality of single exon nucleic acid  
 CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable from the plurality), a single  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC methods of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing  
 CC storage medium which contains a database having a plurality of records  
 CC (each record including data on the expression of a single exon probe  
 CC cited above. The probe, methods and apparatus are useful in gene  
 CC expression analysis. The probes may be used as tools for surveying  
 CC tissues to detect the presence of expressed messages that contain their  
 CC specific exon, or in constructing genome-derived single exon microarrays.  
 CC In addition, the probes are used in identifying and characterizing  
 CC alternative splicing events, in detecting and characterizing gross  
 CC alterations in the genomic locus that includes their exon, in assessing  
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
 CC or in expressing the ORF-encoded peptide. The present sequence is a human  
 CC single exon probe of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
 XX SQ Sequence 211 BP; 32 A; 85 C; 55 G; 39 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3,156-39 Length: 211  
 Score: 367.00 Matches: 70  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 57.43% Indels: 0  
 DB: 12 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x ACH93386 (1-211)

QY 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgGlyTyrArgLeuProSerIle 73  
 Db 1 CAGGACGACCAAGCTACATTTCCCGTGGATTCACCGCGCTACCGCTGCCAAC 60  
 QY 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 93  
 Db 61 GTGGACCAAGTGGCGCTCTCTTGTCTCCCTGTGCGATGCGATGCTGACCTTCTGTGGC 120  
 QY 94 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113  
 Db 121 CCCAAGATCCAGACTGGCTGGATGCCACCCAGCGCGAGCGAGCATCCCGTGTCTGGCG 180  
 QY 114 GluGluLysProThrSerAlaProSerSer 123  
 Db 181 CAGGAGAAGCCACCTCGGCTCCCTCGTCC 210

## RESULT 7

ACH79686  
 ID ACH79686 standard; DNA; 573 BP.  
 XX ACH79686;  
 XX 29-JUL-2004 (first entry)

Human genome derived single exon probe #12881.

Human; probe; ss; Gene expression; single exon probe; microarray;  
 alternative splicing event; genomic alteration.

Homo sapiens.

US2003194704-A1.

16-OCT-2003.

03-APR-2002; 2002US-00029386.

03-APR-2002; 2002US-00029386.

(PENN/) PENN S G.

(RANK/) RANK D R.

(HANZ/) HANZEL D K.

Penn SG, Rank DR, Hanzel DK;

WPI; 2004-119264/12.

New human genome-derived single exon nucleic acid probes useful for human  
 gene expression analysis, for identifying or characterizing alternative  
 splicing events, for assessing genomic alterations or as tools for  
 surveying tissues.

Claim 15; SEQ ID NO 12881; 80pp; English.

The invention relates to a nucleic acid probe for measuring human gene  
 expression, comprising any of the 27,400 fully defined nucleotide  
 sequences in the specification, or their complements or fragments, and  
 encoding at least 8 amino acids of any of the 6888 amino acid sequences  
 fully defined in the specification. The probe is a single exon probe that  
 hybridises under high stringency conditions to a nucleic acid molecule  
 expressed in human cells or tissues. Also included are a spatially-  
 addressable set of single exon nucleic acid probes for measuring human  
 gene expression (comprising a plurality of single exon nucleic acid  
 probes cited above, where each of the plurality of probes is separately  
 and addressably isolatable or amplifiable from the plurality), a single  
 exon microarray for measuring human gene expression, a method of  
 measuring human gene expression, a vector comprising the single exon  
 probe cited above, an ORF-encoded peptide comprising at least 8  
 contiguous amino acids of any of the above-mentioned amino acid  
 sequences (optionally with conservative amino acid substitutions), an  
 isolated antibody that binds specifically to a peptide cited above,  
 methods of selling and/or licensing single exon probes or microarrays to  
 a customer desiring to measure gene expression, a method of providing

human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above). The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030194704

Sequence 573 BP; 86 A; 202 C; 171 G; 114 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 1.33e-38 Length: 573  
Score: 367.00 Matches: 70  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 57.43% Indels: 0  
DB: 12 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x ACH79686 (1-573)

Qy 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerArg 73  
Db 203 CAGGACGACACGGCTACATTCCCGTGGATTCACACCGCGCTACCGCTCGGCCAAC 262  
Qy 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 93  
Db 263 GTGGACCACTGGCCCTCTTGTCTCTCTGCTGCGGATGGCATGCTGACCTTGTGGC 322  
Qy 94 ProTylleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113  
Db 323 CCCAAGATCCAGACTGGCTGGATGCCACCCACGCGGAGCGAGCATCCCGTGTGGGG 382  
Qy 114 GluGluLysProThrSerAlaProSerSer 123  
Db 383 GAGGAGAGCCACCTCGGCTCCCTCGTCC 412

#### RESULT 8

ADBS2642  
ID ADBS2642 standard; DNA; 528 BP.

AC ADBS2642;

DT 04-DEC-2003 (first entry)

DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3184.  
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;  
KW toxicity marker; toxicity progression; drug screening;  
KW primary rat hepatocyte toxicity modelling; gene; ds.

OS Rattus norvegicus.

XX WO2003065993-A2.

XX 14-AUG-2003.

XX 04-FEB-2003; 2003WO-US003482.

XX 04-FEB-2002; 2002US-0353171P.

XX 13-MAR-2002; 2002US-0363534P.

XX 08-APR-2002; 2002US-0370248P.

XX 10-APR-2002; 2002US-0371134P.

XX 10-APR-2002; 2002US-0371135P.

XX 10-APR-2002; 2002US-0371150P.

PR 11-APR-2002; 2002US-0371413P.  
PR 19-APR-2002; 2002US-0373601P.  
PR 19-APR-2002; 2002US-0373602P.  
PR 22-APR-2002; 2002US-0374139P.  
PR 08-MAY-2002; 2002US-0378370P.  
PR 09-MAY-2002; 2002US-0378652P.  
PR 09-MAY-2002; 2002US-0378653P.  
PR 09-MAY-2002; 2002US-0378655P.  
PR 09-JUL-2002; 2002US-0394230P.  
PR 09-JUL-2002; 2002US-0394253P.  
PR 04-SEP-2002; 2002US-0407688P.  
PR 28-JAN-2003; 2003US-0442900P.

XX (GENE-) GENE LOGIC INC.

XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;

PI Elashoff M;

XX WPI; 2003-731472/69.

XX Determining if a compound induces a toxic effect on a tissue or cell, for  
PT identifying hepatotoxic compounds, comprises comparing a gene expression  
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox  
PT mean values.

XX Claim 44; SEQ ID NO 3184; 874pp; English.

XX The present invention describes a method for determining whether a  
CC compound induces a toxic effect on a tissue or cell. The method comprises  
CC preparing a gene expression profile of a tissue or cell sample exposed to  
CC the compound, and comparing the gene expression profile to a database  
CC comprising data or information on the Tox mean and non-Tox mean value.  
CC The method is useful for predicting or identifying at least one toxic  
CC effect, particularly hepatotoxicity, of a test or unknown compound. The  
CC genes listed in the specification are useful as diagnostic or toxicity  
CC markers for the prediction or identification of the physiological state  
CC of tissue or cell sample that has been exposed to a compound, or to  
CC identify or predict the toxic effects of a compound or an agent. These  
CC may also be used as markers for monitoring toxicity progression or for  
CC drug screening. The present sequence represents a primary rat hepatocyte  
CC toxicity modelling related gene sequence from the present invention.

SQ Sequence 528 BP; 114 A; 170 C; 128 G; 116 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 7.32e-38 Length: 528  
Score: 361.00 Matches: 67  
Percent Similarity: 79.49% Conservative: 26  
Best Local Similarity: 57.26% Mismatches: 18  
Query Match: 56.49% Indels: 6  
DB: 10 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x ADBS2642 (1-528)

Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
Db 181 ATTACACATCGGCTCTCAGACATCGTATGGAGAGGACAGGTTCTCTGTGAACCTGGAC 240  
Qy 27 ValIleHisPheSerProGluAspLeuThrValIleValGlnAspAspPheValGluIle 46  
Db 241 GTGAAGCACTTCTCTCCAGAGAACTCAAGTCAAGGTTCTGGGAGAGCTGATGAGGTG 300  
Qy 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
Db 301 CACGGCAGCAGCAGAGAGCGCAGGACGAACATGGTTCATCTCCAGGAGTTCACAGG 360  
Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 361 AAGTACCGGATCCAGCGGACGTGGATCTCTCACCATTACTTCTTCCCTGTCTCGGAT 420  
Qy 87 GlyMetLeuThrPheCysGlyProIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
Db 421 GGAGTCTCTACTGTGAATGGACCAAGGAACAG-----GCCTCTGGCCCTGAG 468







Alignment Scores:  
 Pred. No.: 2.52e-37 Length: 1247  
 Score: 361.00 Matches: 67  
 Percent Similarity: 79.49% Conservative: 26  
 Best Local Similarity: 57.26% Mismatches: 18  
 Query Match: 56.49% Indels: 6  
 DB: 12 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x ADP72693 (1-1247)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 763 ATTGACACTGGCTCTCAGACATCGTATGAGAGGACACAGGTTCTCTGGAACCTGGAC 822  
 QY 27 VallyshHisPheSerProGluAspLeuThrValIysValGlnAspPheValGluIle 46  
 Db 823 GTGAAGCACCTCTCCAGAGGAACCTCAAAGTCAAGGTTCTGGAGACGTGATGAGGTG 882  
 QY 47 HisGlyIysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
 Db 883 CACGCAAGCACGAAGAGCGCCAGGACGAACATGGCTTCATCTCCAGGAGTTCCACAGG 942  
 QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 943 AAGTACCGGATCCAGCCGAGCGTGGATCTCTCACCATTACTTCTTCCCTGTCATCGGAT 1002  
 QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 1003 GGAGTCTCTCACTGTGAATGACCAAGGAACAG-----GCCCTGGCCCTGAG 1050  
 QY 107 ArgAlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
 Db 1051 CGCACCATTCCTCCATCCCGGTGAAGAGAGCGCTGCTGCTCACTGCACGCCCT 1101

RESULT 12  
 ABZ35138  
 ID ABZ35138 standard; cDNA; 548 BP.  
 AC ABZ35138;  
 XX  
 XX  
 DT 05-FEB-2003 (first entry)  
 XX  
 DE Human gene expression profile polynucleotide SEQ ID NO 250.  
 XX  
 KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;  
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;  
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;  
 KW gene expression; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200274979-A2.  
 XX  
 PN 26-SEP-2002.  
 XX  
 PF 20-MAR-2002; 2002WO-US008456.  
 XX  
 PR 20-MAR-2001; 2001US-0276947P.  
 XX  
 XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
 PA  
 PI Wan J, Wang Y;  
 XX  
 DR WPI; 2002-740862/80.  
 XX  
 PT New gene expression profile generated from primary, endothelial,  
 PT epithelial, and muscle cell types, useful for identifying disease  
 PT pathologies involving alterations of gene expression, e.g. cancer.  
 XX  
 PS Claim 10; Page 429; 850pp; English.  
 XX  
 CC The invention relates to a gene expression profile comprising one or more  
 CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type

CC is a coronary artery endothelium, umbilical artery or vein endothelium,  
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery  
 CC endothelium, myometrium microvascular endothelium, keratinocyte  
 CC epithelium, bronchial epithelium, mammary epithelium, prostate  
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,  
 CC small airway epithelium, renal epithelium, umbilical artery smooth  
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,  
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,  
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,  
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,  
 CC osteoblasts or prostate stromal cell. The gene expression profile is used  
 CC for determining the level of RNA expression for a sample, determining the  
 CC phenotype of a cell and distinguishing cell types. The gene or a protein  
 CC expression profile is useful in identifying disease pathologies involving  
 CC alterations of gene expression. The assessment of expression profiles may  
 CC provide meaningful information with respect to tumour type and stage,  
 CC treatment methods, and prognosis. The gene or protein expression profile  
 CC may also be used for creating microarrays. The microarray is useful for  
 CC genetic and physical mapping of genomes, DNA sequencing, genetic or  
 CC medical diagnosis, genotyping of organisms, confirming cell or tissue  
 CC identifications and in identifying promising antibiotics, antiviral or  
 CC antifungal agents  
 CC  
 SQ Sequence 548 BP; 120 A; 178 C; 127 G; 123 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.92e-37 Length: 548  
 Score: 358.00 Matches: 67  
 Percent Similarity: 78.63% Conservative: 25  
 Best Local Similarity: 57.26% Mismatches: 19  
 Query Match: 56.03% Indels: 6  
 DB: 6 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x ABZ35138 (1-548)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 201 ATTGACACTGGCTCTCAGACATCGCTGGAGAGGACAGGTTCTCTGCAACCTGGAT 260  
 QY 27 VallyshHisPheSerProGluAspLeuThrValIysValGlnAspPheValGluIle 46  
 Db 261 GTGAAGCACCTCTCTCCAGAGGAGCTCAAGTCAAGTGTGGTGTGATGTGATGAGTG 320  
 QY 47 HisGlyIysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
 Db 321 CACGCAACATCATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGAGTTCCACAGG 380  
 QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 381 AAATACCGGATCCAGCTGATGTGGACCTCTCACCATTACTTCACTCCCTGTCATCTGAT 440  
 QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 441 GGGGTCTCTCACTGTGAATGACCAAGGAAGCAA-----GCCCTGGCCAGAG 488  
 QY 107 ArgAlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
 Db 489 CGCACCATTCCTCCATCCCGGTGAAGAGAGCGCTGCTGCTCACTGCACGCCCTCC 539

RESULT 13  
 ADR46388  
 ID ADR46388 standard; DNA; 528 BP.  
 XX  
 AC ADR46388;  
 XX  
 DT 04-NOV-2004 (first entry)  
 XX  
 DE Human lens crystalline protein coding sequence.  
 XX  
 KW cytostatic; immunotoxin; cancer; mitochondrial malate dehydrogenase;  
 KW enzyme; human; MDH; ds; gene; lens crystalline protein.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 1..528  
 FT /\*tag= a  
 FT /product= "lens crystalline protein"  
 XX WO2004070012-A2.  
 XX 19-AUG-2004.  
 XX 02-FEB-2004; 2004WO-US002974.  
 XX 02-FEB-2003; 2003US-00444191.  
 PR 03-FEB-2003; 2003US-0444191P.  
 PR 08-APR-2003; 2003US-0460855P.  
 XX (PALO-) PALO ALTO INST MOLECULAR MEDICINE.  
 PA Wright SC, Larrick JW, Nock SR, Wilson DS;  
 PI WPI; 2004-604434/58.  
 DR P-PSDB; ADR46387.  
 XX New compositions comprising proteins and encoding nucleic acids having a  
 PT DNA nuclease or cell killing activity and are operably linked to cancer  
 PT cell binding antibodies or growth factors, useful for treating cancer.  
 XX  
 PS Disclosure; SEQ ID NO 35; 225pp; English.  
 XX  
 CC The present invention relates to a composition comprising an isolated  
 CC amino acid sequence that comprises a portion of human mitochondrial  
 CC malate dehydrogenase protein (MDH), particularly the minimum activator of  
 CC DNA fragmentation and activator of DNA fragmentation sequences. The  
 CC composition is useful for treating cancer chosen from liver cancer,  
 CC gastric cancer, head cancer, neck cancer, lung cancer, breast cancer,  
 CC prostate cancer, cervical cancer, pancreatic cancer, colon cancer,  
 CC ovarian cancer, stomach cancer, oesophagus cancer, mouth cancer, tongue  
 CC cancer, gum cancer, skin cancer, muscle cancer, heart cancer, bronchial  
 CC cancer, cartilage cancer, bone cancer, testis cancer, kidney cancer,  
 CC endometrium cancer, uterus cancer, bladder cancer, bone marrow cancer,  
 CC lymphoma cancer, spleen cancer, thymus cancer, thyroid cancer, brain  
 CC cancer, neuron cancer, gall bladder cancer, ocular cancer, joint cancer,  
 CC glioblastoma, mesothelioma, lymphoma, leukaemia, melanoma, squamous cell  
 CC carcinoma, osteosarcoma, and Kaposi's sarcoma. The present sequence is a  
 CC coding sequence shown in the exemplification of the invention.  
 XX  
 SQ Sequence 528 BP; 115 A; 167 C; 123 G; 123 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 6.16e-37 Length: 528  
 Score: 354.00 Matches: 66  
 Percent Similarity: 78.45% Conservative: 25  
 Best Local Similarity: 56.90% Mismatches: 19  
 Query Match: 55.40% Indels: 6  
 DB: 13 Gaps: 2  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x ADR46388 (1-528)  
 QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27  
 DB 184 GACACTGGACTCTCAGATGCGCTCGGAGAGAGACAGGTTCTCTGTAACCTGGATGTG 243  
 QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 47  
 DB 244 AAGCACTTCTCCCGAGAGAACTCAAGTTAAGGTGTTGGGAGATGTAATGAGTGCAT 303  
 QY 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 67  
 DB 304 GGAAACATGAGAGCGCCAGCATGACATGTTTCATCTCCAGGAGTTCACAGGAA 363  
 QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87  
 DB 364 TACCGATCCGACTGATGATAGACCCCTCTCACCAATTACTTCTCATCTGTCATGGG 423

QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107  
 DB 424 GTCCTCACTGTGAATGACCAAGGAACAG-----GTCTCTGGCCCTGAGCGC 471  
 QY 108 AlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121  
 DB 472 ACCATTCCCATCCCGTGAAGAGAGAGCTGTCTGTACCCGAGCCCC 519

RESULT 14  
 ADE75374  
 ID ADE75374 standard; DNA; 537 BP.  
 XX  
 AC ADE75374;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human mutant alphaB-crystallin fragment-encoding DNA.  
 XX  
 KW Hybrid protein chaperone; protein stabilisation; heat shock protein;  
 KW SHSP family; protein aggregation inhibition; cell death inhibition;  
 KW genome stability pathway inhibition; protein denaturation identification;  
 KW protein conformation related disease; cardiomyopathy; cataract;  
 KW neurodegenerative disease; cardiac; ophthalmological; neuroprotective;  
 KW gene therapy; alphaB-crystallin; wild-type; human; cassette mutagenesis;  
 KW mutant; ds.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_feature /tag= a  
 FT mutation /note= "Unique Aval site for cassette mutagenesis"  
 FT replace(497,T)  
 FT /tag= b  
 FT /note= "Mutagenesis to G from T at this site generates a  
 FT unique Aval site between bases 493-498"  
 XX WO2003091266-A2.  
 PD 06-NOV-2003.  
 XX  
 PF 23-APR-2003; 2003WO-GB001721.  
 XX  
 PR 23-APR-2002; 2002GB-00009334.  
 XX  
 PA (UVDU-) UNIV DUNDEE.  
 XX  
 PI Quinlan R;  
 XX  
 WI WPI; 2003-865571/80.  
 XX  
 PT New hybrid protein chaperone (e.g. heat shock protein) useful for  
 PT stabilizing proteins and/or protein activities, or as an agent to prevent  
 PT protein aggregation, or for treating diseases involving altered protein  
 PT conformations.  
 XX  
 PS Disclosure; Fig 12; 45pp; English.  
 XX  
 CC The invention relates to a hybrid protein chaperone for stabilising  
 CC proteins and/or protein activities. Protein chaperones (also known as  
 CC heat shock proteins) are divided into 4 families on the basis of their  
 CC primary sequence and chaperone properties: HSP90, HSP70, HSP60 and SHSP  
 CC (small heat shock protein). The invention is based upon the finding that  
 CC among the SHSP family, which have a general structure of a central domain  
 CC (called the alpha-crystallin domain) flanked by N and C-terminal regions,  
 CC replacement of one or more regions of an SHSP with the corresponding  
 CC region from a second SHSP can improve the activity compared to native  
 CC SHSPs. In a particular embodiment of the invention, the hybrid chaperone  
 CC is a hybrid SHSP designated alphaB-HSP27 comprising the N-terminus and  
 CC central portion of alphaB-crystallin and the C-terminal tail of HSP27.  
 CC However, the hybrid protein chaperones of the invention can comprise



```
DB: 10 Gaps: 2
US-10-657-740-1_COPY_51_173 (1-123) x ADE75375 (1-537)
QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
D 192 GACACTGGACTCTCAGAGATGCGCTGGAGAGGACAGGTTCTCTGTCACCTGGATGG 251
QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluLeuHis 47
D 252 AAGCACTTCTCCCGCAGAGAACTCAAGTAAAGTGTGGAGAGTGTGATTGAGGTGCAT 311
QY 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 67
D 312 GAAACATGATGAAGAGCGCCAGATGAACATGTTTCTCTCAGGAGTTCCACAGGAAA 371
QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
D 372 TACCGATCCAGCTGATGAGCCCTCTCACCATTACTTCATCCTCTGTCATCTGATGG 431
QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
D 432 GTCCTCACTGTGATGACCAAGGAACAG-----GTCTCTGCCCTGAGCGC 479
QY 108 AlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121
D 480 ACCATTCCCATCACCGGTGAAGAGAGCGTGTGTCTCAGCGAGCCCCC 527

RESULT 16
ID AAX39668 standard; DNA; 691 BP.
XX
AC AAX39668;
XX
DT 02-JUL-1999 (first entry)
XX
DE Renal cancer associated gene.
XX
KW Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer; ss.
XX
OS Homo sapiens.
XX
FN WO9904265-A2.
XX
PD 28-JAN-1999.
XX
PF 15-JUL-1998; 98WO-US014679.
XX
PR 17-JUL-1997; 97US-00896164.
PR 10-OCT-1997; 97US-0061599P.
PR 10-OCT-1997; 97US-0061765P.
PR 10-OCT-1997; 97US-00948705.
PR 11-OCT-1997; 97GB-00021697.
PR 22-JUN-1998; 98US-00102322.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
XX Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;
PI O'hare M, Obata Y, Pfeundschoh M, Tureci O, Sahin U;
XX
XX WPI; 1999-132448/11.
XX
XX New isolated cancer associated nucleic acids and polypeptides - isolated
PT using sera from cancer patients, used to develop products for the
PT diagnosis, monitoring or treatment of cancers.
XX
PS Claim 67; Page 478; 787pp; English.
XX
XX The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
```

```
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer
XX
```

SQ Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	9,07e-37	Length:	691
Score:	354.00	Matches:	66
Percent Similarity:	78.45%	Conservative:	25
Best Local Similarity:	56.90%	Mismatches:	19
Query Match:	55.40%	Indels:	6
DB:	2	Gaps:	2

US-10-657-740-1\_COPY\_51\_173 (1-123) x AAX39668 (1-691)

```
QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
D 209 GACACTGGACTCTCAGAGATGCGCTGGAGAGGACAGGTTCTCTGTCAACCTGGATGG 268
QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluLeuHis 47
D 269 AAGCACTTCTCCCGCAGAGAACTCAAGTAAAGTGTGGAGAGTGTGATTGAGGTGCAT 328
QY 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 67
D 329 GAAACATGATGAAGAGCGCCAGGATGAACATGTTTCTCCTCAGGAGTTCCACAGGAAA 388
QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
D 389 TACCGATCCAGCTGATGAGCCCTCTCACCATTACTTCATCCTCTGTCATCTGATGG 448
QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
D 449 GTCCTCACTGTGATGACCAAGGAACAG-----GTCTCTGCCCTGAGCGC 496
QY 108 AlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121
D 497 ACCATTCCCATCACCGGTGAAGAGAGCGTGTGTCTCAGCGAGCCCCC 544
```

#### RESULT 17

ABL65204  
ID ABL65204 standard; DNA; 691 BP.

XX ABL65204;

XX 15-MAY-2002 (first entry)

XX Lung cancer related gene sequence SEQ ID NO:3541.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

PR 05-JUN-2000; 2000US-0209531P.

PR	18-SEP-2000;	2000US-02331133P.
PR	18-SEP-2000;	2000US-02331136P.
PR	20-SEP-2000;	2000US-0234009P.
PR	20-SEP-2000;	2000US-0234034P.
PR	20-SEP-2000;	2000US-0234052P.
PR	22-SEP-2000;	2000US-0234509P.
PR	22-SEP-2000;	2000US-0234567P.
PR	25-SEP-2000;	2000US-0234923P.
PR	25-SEP-2000;	2000US-0234924P.
PR	25-SEP-2000;	2000US-0235077P.
PR	25-SEP-2000;	2000US-0235082P.
PR	25-SEP-2000;	2000US-0235134P.
PR	26-SEP-2000;	2000US-0235280P.
PR	26-SEP-2000;	2000US-0235637P.
PR	26-SEP-2000;	2000US-0235638P.
PR	27-SEP-2000;	2000US-0235711P.
PR	27-SEP-2000;	2000US-0235720P.
PR	27-SEP-2000;	2000US-0235840P.
PR	28-SEP-2000;	2000US-0235863P.
PR	28-SEP-2000;	2000US-0236032P.
PR	28-SEP-2000;	2000US-0236033P.
PR	28-SEP-2000;	2000US-0236034P.
PR	28-SEP-2000;	2000US-0236109P.
PR	28-SEP-2000;	2000US-0236111P.
PR	29-SEP-2000;	2000US-0236842P.
PR	29-SEP-2000;	2000US-0236891P.
PR	02-OCT-2000;	2000US-0237172P.
PR	02-OCT-2000;	2000US-0237173P.
PR	02-OCT-2000;	2000US-0237278P.
PR	02-OCT-2000;	2000US-0237294P.
PR	02-OCT-2000;	2000US-0237295P.
PR	03-OCT-2000;	2000US-0237316P.
PR	03-OCT-2000;	2000US-0237425P.
PR	03-OCT-2000;	2000US-0237598P.
PR	03-OCT-2000;	2000US-0237604P.
PR	03-OCT-2000;	2000US-0237606P.
PR	03-OCT-2000;	2000US-0237608P.
PR	01-NOV-2000;	2000US-0244867P.
PR	01-NOV-2000;	2000US-0245084P.
XX	(AVAL-)	AVALON PHARM.
XX	Young PE,	Augustus M,
PI	Soppet DR,	Weaver Z;
XX	WPI;	2002-188264/24.
XX	Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.	
PS	Claim 1;	SEQ ID NO 3541; 44pp; English.
XX	The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences given in ABL61664 to ABL70110, or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour	
XX	Sequence 691 BP;	168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	9.07e-37	Length:	691
Score:	354.00	Matches:	66
Percent Similarity:	78.45%	Conservative:	25
Best Local Similarity:	56.90%	Mismatches:	19
Query Match:	55.40%	Indels:	6
DB:	6	Gaps:	2
US-10-657-740-1_COPY_51_173 (1-123) x ABL65204 (1-691)			
Qy	8	AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal	27
Db	209	GACACTGGACTCTCAGAGATGGCCCTGGAGAGGACAGGTTCTCTGTCAACCTGGATGG	268
Qy	28	LysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIleHis	47
Db	269	AAGCACTTCTCCCCAGAGGAATCAAAGTTAAGGTGTGGAGATCTGATTGAGGTGCAT	328
Qy	48	GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg	67
Db	329	GGAAAAACATGAAGAGCGCCAGATGAACATGTTTTCATCTCCAGGATTCACAGGAAA	388
Qy	68	TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly	87
Db	389	TACCGGATCCCGAGCTGATGTAGACCCCTCTCACCATTACTTCATCCCTGTCTATCTGATGG	448
Qy	88	MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg	107
Db	449	GTCCTCACTGTGATGGACCAAGGAACAG-----GTCTGTGGCCCTGAGCGC	496
Qy	108	AlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro	121
Db	497	ACCATTCCCATCACCCGTGAAGAGAAGCGCTGCTGCACCGCAGCCCCC	544
RESULT 18			
ABL62427			
ID	ABL62427 standard; DNA; 691 BP.		
XX	AC	ABL62427;	
XX	DT	15-MAY-2002 (first entry)	
XX	DE	Colon adenocarcinoma related gene sequence SEQ ID NO:764.	
XX	KW	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;	
XX	KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;	
XX	KW	cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;	
XX	KW	gene; ds.	
XX	XX	Homo sapiens.	
XX	XX	WO200194629-A2.	
XX	PD	13-DEC-2001.	
XX	PF	30-MAY-2001; 2001WO-US010838.	
XX	XX	05-JUN-2000; 2000US-0209473P.	
PR	PR	05-JUN-2000; 2000US-0209531P.	
PR	PR	18-SEP-2000; 2000US-0233133P.	
PR	PR	20-SEP-2000; 2000US-0233617P.	
PR	PR	20-SEP-2000; 2000US-0234009P.	
PR	PR	20-SEP-2000; 2000US-0234034P.	
PR	PR	20-SEP-2000; 2000US-0234052P.	
PR	PR	22-SEP-2000; 2000US-0234509P.	
PR	PR	22-SEP-2000; 2000US-0234567P.	
PR	PR	25-SEP-2000; 2000US-0234923P.	
PR	PR	25-SEP-2000; 2000US-0234924P.	
PR	PR	25-SEP-2000; 2000US-0235077P.	
PR	PR	25-SEP-2000; 2000US-0235082P.	
PR	PR	25-SEP-2000; 2000US-0235134P.	
PR	PR	25-SEP-2000; 2000US-0235280P.	







```

XX OS Homo sapiens.
XX PN WO2003068940-A2.
XX PD 21-AUG-2003.
XX PF 14-FEB-2003; 2003WO-US004594.
XX PR 14-FEB-2002; 2002US-0356911P.
XX PA (CURA-) CURAGEN CORP.
XX PA (HOFF ) HOFFMANN LA ROCHE INC.
XX PI Jackson A, Ooi CE, Lewin DA, Cuthill S;
XX WPI; 2003-689668/65.
XX P-PSDB; ADF09576.
XX
XX New purified complex comprising a first polypeptide and a second
XX polypeptide, useful for identifying agents for treating/preventing a
XX condition involving altered level of the complex e.g. human papilloma
XX virus infection, or cancer.
XX
XX Example 3; SEQ ID NO 187; 156pp; English.
XX
XX The invention relates to a novel purified complex comprising a first
XX polypeptide and a second polypeptide, where the polypeptides comprise
XX defined amino acid sequences listed in the specification, and where the
XX first polypeptide binds to the second polypeptide. A complex of the
XX invention has virucide and cytostatic activity, and may have a use as a
XX vaccine. The complex is useful for identifying agents for treating or
XX preventing a conditions involving altered level of the complex, e.g.
XX human papilloma virus (HPV) infection, or cancer. The compositions,
XX antibodies, vectors and methods are useful for treating such diseases.
XX The sequences shown in ADF09584-ADF09697 represent cDNA's of the
XX invention.
XX
XX SQ Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.07e-37 Length: 691
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservative: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 10 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x ADF09686 (1-691)
Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 209 GACACTGGACTCTCAGAGATGGCGCTGGAGAGGACAGGTTCTCTGTCAACCTGGATGTG 268
Qy 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLeHis 47
Db 269 AAGCACTTCTCCCGAGAGAACTCAAAGTTAAGGTGTGGAGATGTGATTGATGTGCAT 328
Qy 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgArg 67
Db 329 GGAATAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCACAGGAAA 388
Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
Db 389 TACCGATCCAGCTGATAGACCTCTCACCATTACTTCTCATCTCATCTGATGGG 448
Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 449 GTCTCTACTGTAATGACCAAGAAACAG-----GTCTCTGGCCCTGAGCGC 496
Qy 108 AlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121
Db 497 ACCATTCCCATCCACCCGTGAAGAGAGCGCTCTCTCTCACCAGCCGCC 544

```

```

RESULT 22
ADN05800
ID ADN05800 standard; cDNA; 691 BP.
XX AC ADN05800;
XX DT 01-JUL-2004 (first entry)
XX DE Antipsoriatic cDNA sequence #1131.
XX KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX PN WO2004028479-A2.
XX PD 08-APR-2004.
XX PF 25-SEP-2003; 2003WO-US030907.
XX PR 25-SEP-2002; 2002US-0414006P.
XX PA (GETH ) GENENTECH INC.
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX WPI; 2004-305105/28.
XX P-PSDB; ADN05801.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
XX
XX Claim 1; SEQ ID NO 2194; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polynucleotides of the invention.
XX
XX SQ Sequence 691 BP; 168 A; 208 C; 148 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.07e-37 Length: 691
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservative: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 12 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x ADN05800 (1-691)
Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 209 GACACTGGACTCTCAGAGATGGCGCTGGAGAGGACAGGTTCTCTGTCAACCTGGATGTG 268
Qy 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLeHis 47
Db 269 AAGCACTTCTCCCGAGAGAACTCAAAGTTAAGGTGTGGAGATGTGATTGATGTGCAT 328
Qy 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgArg 67
Db 329 GGAATAACATGAAGAGCGCCAGGATGAACATGGTTTCATCTCCAGGGAGTTCACAGGAAA 388
Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
Db 389 TACCGATCCAGCTGATAGACCTCTCACCATTACTTCTCATCTCATCTGATGGG 448
Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107

```



Query Match: 55.40% Indels: 6  
DB: 3 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x AAC03893 (1-856)

Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspIlyPheValIlePheLeuAspVal 27  
412 GACACTGGACTCTCAGAGATCGGCTGGAGAGACAGGTTCTCTGTCAACCTGGATGTG 471

Db 28 LysHisPheSerProGluAspLeuThrValIlyValGlnAspAspPheValGluIleHis 47  
472 AAGCACTTCTCCCGAGAGAACTCAAGTTAAGGTGTGGAGATGTGATTGAGTGCAT 531

Qy 48 GlyIlyHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 67  
532 GGAACAACATGAAGAGCGCCAGATGAACATGTTTCATCTCCAGGAGTTCACAGGAAA 591

Db 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87  
592 TACCGATCCAGCTGATGTAGACCTCTCACCATTTACTTCTCTGTCTATCTGTATGGG 651

Qy 88 MetLeuThrPheCysGlyProIlyIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107  
652 GTCCTCACTGTGAATGGACCAAGAAACAG-----GTCCTCTGGCCCTGAGCGC 699

Db 108 AlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
700 ACCATTCCCATCACCCTGTGAAGAGAGCCTCTGTCTCACCAGCAGCCCC 747

RESULT 25  
AAC10867  
ID AAC10867 standard; cDNA; 893 BP.  
XX AC AAC10867;  
XX DT 06-OCT-2000 (first entry)  
XX DE Human secreted protein 5' EST, SEQ ID NO: 14942.  
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX KW gene therapy; chromosome mapping; ss.  
XX OS Homo sapiens.  
XX FN EP1033401-A2.  
XX PD 06-SEP-2000.  
XX PF 21-FEB-2000; 2000EP-00200610.  
XX PR 26-FEB-1999; 99US-0122487P.  
XX PA (GEST ) GENSET.  
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX DR WPI; 2000-500381/45.  
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX PS Claim 1; SEQ ID NO 14942; 71pp + Sequence Listing; English.  
XX CC The present sequence is one of a large number of 5' ESTs derived from  
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively  
XX identified within the present sequence. The 5' ESTs were prepared from  
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
XX sequences usually correspond mainly to the 3' untranslated region (UTR)  
XX of the mRNA because they are often obtained from oligo-dT primed cDNA  
XX libraries. Such ESTs are not well suited for isolating cDNA sequences  
XX derived from the 5' ends of mRNAs and even in those cases where longer  
XX cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'

CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors  
XX Sequence 893 BP; 221 A; 259 C; 215 G; 198 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.31e-36 Length: 893  
Score: 354.00 Matches: 66  
Percent Similarity: 78.45% Conservative: 25  
Best Local Similarity: 56.90% Mismatches: 19  
Query Match: 55.40%  
DB: 3 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x AAC10867 (1-893)

Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspIlyPheValIlePheLeuAspVal 27  
449 GACACTGGACTCTCAGAGATCGGCTGGAGAGACAGGTTCTCTGTCAACCTGGATGTG 508

Db 28 LysHisPheSerProGluAspLeuThrValIlyValGlnAspAspPheValGluIleHis 47  
509 AAGCACTTCTCCCGAGAGAACTCAAGTTAAGGTGTGGAGATGTGATTGAGTGCAT 568

Qy 48 GlyIlyHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 67  
569 GGAACAACATGAAGAGCGCCAGATGAACATGTTTCATCTCCAGGAGTTCACAGGAAA 628

Db 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87  
629 TACCGATCCAGCTGATGTAGACCTCTCACCATTTACTTCTCACCCTGTCTATCTGTATGGG 688

Qy 88 MetLeuThrPheCysGlyProIlyIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107  
689 GTCCTCACTGTGAATGGACCAAGAAACAG-----GTCCTCTGGCCCTGAGCGC 736

Qy 108 AlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
737 ACCATTCCCATCACCCTGTGAAGAGAGCCTCTGTCTCACCAGCAGCCCC 784

RESULT 26  
AAC10863  
ID AAC10863 standard; cDNA; 911 BP.  
XX AC AAC10863;  
XX DT 06-OCT-2000 (first entry)  
XX DE Human secreted protein 5' EST, SEQ ID NO: 14938.  
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX KW gene therapy; chromosome mapping; ss.  
XX OS Homo sapiens.  
XX FN EP1033401-A2.  
XX PD 06-SEP-2000.  
XX PF 21-FEB-2000; 2000EP-00200610.  
XX PR 26-FEB-1999; 99US-0122487P.  
XX PA (GEST ) GENSET.  
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX DR WPI; 2000-500381/45.  
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for

```
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX Claim 1; SEQ ID NO 14938; 71pp + Sequence Listing; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
XX ESTs are derived from mRNAs with intact 5' ends and can therefore be used
XX to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors
XX
XX Sequence 911 BP; 221 A; 260 C; 211 G; 219 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.35e-36 Length: 911
XX Score: 354.00 Matches: 66
XX Percent Similarity: 78.45% Conservative: 25
XX Best Local Similarity: 56.90% Mismatches: 19
XX Query Match: 55.40% Indels: 6
XX DB: 3 Gaps: 2
XX
XX US-10-657-740-1_COPY_51_173 (1-123) x AAC10863 (1-911)
XX
XX QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
XX Db 467 GACACTGGACTCTCAGAGATGGCTGGAGAGGACAGGTTCTGTCAACCTGGATGTG 526
XX
XX QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluLeHis 47
XX Db 527 AAGCACTTCTCCCGACAGAACTCAAGTAAAGTGTGGGAGATGTGATTGAGTGCAT 586
XX
XX QY 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgArg 67
XX Db 587 GGAACAACTGAAGAGCGCCAGGATGACATGGTTTTCATCTCCAGGGAGTTCCACAGGAAA 646
XX
XX QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
XX Db 647 TACCGATCCAGCTGATGTAGACCTCTCACCATTACTTCTCCTCTCATCTGTATGGG 706
XX
XX QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
XX Db 707 GTCTCTCACTGTGAATGGACCAAGGAACAG-----GTCTCTGGCCCTGAGGCG 754
XX
XX QY 108 AlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121
XX Db 755 ACCATTCCATCACCCTGGAAGAGAGCCTGTGTGTACCGAGCCCC 802
XX
XX RESULT 27
XX AAC10862
XX ID AAC10862 standard; cDNA; 913 BP.
XX XX
XX AC AAC10862;
XX
XX DT 06-OCT-2000 (first entry)
XX
XX DE Human secreted protein 5' EST, SEQ ID NO: 14937.
XX
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX OS Homo sapiens.
XX
XX PN EP1033401-A2.
XX
XX PD 06-SEP-2000.
```

```
XX
XX 21-FEB-2000; 2000EP-00200610.
XX
XX 26-FEB-1999; 99US-0122487P.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 14937; 71pp + Sequence Listing; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
XX ESTs are derived from mRNAs with intact 5' ends and can therefore be used
XX to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors
XX
XX Sequence 913 BP; 228 A; 257 C; 220 G; 202 T; 0 U; 6 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.35e-36 Length: 913
XX Score: 354.00 Matches: 66
XX Percent Similarity: 78.45% Conservative: 25
XX Best Local Similarity: 56.90% Mismatches: 19
XX Query Match: 55.40% Indels: 6
XX DB: 3 Gaps: 2
XX
XX US-10-657-740-1_COPY_51_173 (1-123) x AAC10862 (1-913)
XX
XX QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
XX Db 469 GACACTGGACTCTCAGAGATGGCTGGAGAGGACAGGTTCTGTCAACCTGGATGTG 528
XX
XX QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluLeHis 47
XX Db 529 AAGCACTTCTCCCGACAGAACTCAAGTAAAGTGTGGGAGATGTGATTGAGTGCAT 588
XX
XX QY 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgArg 67
XX Db 589 GGAACAACTGAAGAGCGCCAGGATGACATGGTTTTCATCTCCAGGGAGTTCCACAGGAAA 648
XX
XX QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
XX Db 649 TACCGATCCAGCTGATGTAGACCTCTCACCATTACTTCTCCTCTCATCTGTATGGG 708
XX
XX QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
XX Db 709 GTCTCTCACTGTGAATGGACCAAGGAACAG-----GTCTCTGGCCCTGAGGCG 756
XX
XX QY 108 AlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121
XX Db 757 ACCATTCCATCACCCTGGAAGAGAGCCTGTGTGTACCGAGCCCC 804
XX
XX RESULT 28
XX AAC10861
XX ID AAC10861 standard; cDNA; 927 BP.
XX
XX AC AAC10861;
```

```
XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 14936.
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN EP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-00200610.
XX PR 26-FEB-1999; 99US-0122487P.
XX PA (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX PS Claim 1; SEQ ID NO 14936; 71pp + Sequence Listing; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX CC identified within the present sequence. The 5' ESTs were prepared from
XX CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX CC sequences usually correspond mainly to the 3' untranslated region (UTR)
XX CC of the mRNA because they are often obtained from oligo-dT primed cDNA
XX CC libraries. Such ESTs are not well suited for isolating cDNA sequences
XX CC derived from the 5' ends of mRNAs and even in those cases where longer
XX CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
XX CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
XX CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
XX CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX CC They are used to obtain upstream regulatory sequences and to design
XX CC expression and secretion vectors
XX SQ Sequence 927 BP; 230 A; 267 C; 224 G; 206 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.38e-36 Length: 927
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservatave: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 3 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x AAC10861 (1-927)

Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 483 GACACTGGACTCTCAGAGATCGGCTGGAGAGACAGGTTCTCTGTCAACTGGATGTG 542

Qy 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 47
Db 543 AAGCACTTCTCCACAGGAACTCAAGTTAAGGTGTGGAGATGTGATTGAGGTGCAT 602

Qy 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgArg 67
Db 603 GGAACATGATGAAGCGCCAGGATCAACATGTTTCATCTCCAGGAGTTCACAGGAAA 662

Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
Db 663 TACCGGATCCAGCTGATGTAGACCTCTACCATTTACTTCATCCCTGTCATCTGATGGG 722
```

```
Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 723 GTCCTCACTGTGAATGGACCAAGAAACAG-----GTCTGTGGCCCTGAGCGC 770

Qy 108 AlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121
Db 771 ACCATTCCATCACCCTGGAGAGAGCGCTCTGTCCAGCGAGCCCC 818

RESULT 29
AAC03892
ID AAC03892 standard; cDNA; 942 BP.
XX AC AAC03892;
XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 3890.
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN EP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-00200610.
XX PR 26-FEB-1999; 99US-0122487P.
XX PA (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX DR P-FSDB; AAG03886.
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX PS Claim 1; SEQ ID NO 3890; 71pp + Sequence Listing; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. An ORF has been identified within the
XX CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX CC derived from 30 different tissues. EST sequences usually correspond
XX CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX CC well suited for isolating cDNA sequences derived from the 5' ends of
XX CC mRNAs and even in those cases where longer cDNA sequences have been
XX CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX CC mRNAs with intact 5' ends and can therefore be used to obtain full length
XX CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX CC gene therapy and chromosome mapping procedures. They are used to obtain
XX CC upstream regulatory sequences and to design expression and secretion
XX CC vectors
XX SQ Sequence 942 BP; 223 A; 267 C; 224 G; 228 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.42e-36 Length: 942
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservatave: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 3 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x AAC03892 (1-942)

Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 723 GTCCTCACTGTGAATGGACCAAGAAACAG-----GTCTGTGGCCCTGAGCGC 770
```



Db 498 GACACTGGACTCTCAGAGATCGCGCTGGAGAACAGACAGGTTCTCTGTCAACCTGGATGTG 557  
 Qy 28 LysHisPheSerProGluAspLeuThrValIysValGlnAspAspPheValGluIleHis 47  
 Db 558 AAGACACTTCTCCCGAGGAACTCAAGTTAAGGTGTGGAGATGTGATTGAGGTGCAT 617  
 Qy 48 GlyIysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 67  
 Db 618 GGAACATGAAGAGCCAGGATGAACATGTTTCATCTCCAGGAGTTCACAGAAA 677  
 Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87  
 Db 678 TACCGATCCAGCTGATGTAGACCTCTCACCATTTCTATCTCCCTGTCTCATCTGATGG 737  
 Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107  
 Db 738 GTCTCTACTGTGAATGACCAAGAAACAG-----GTCTCTGGCCCTGAGCGC 785  
 Qy 108 AlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
 Db 786 ACCATTCCCATCACCCTGGAAGAGAGCTGTCTGTCAACCGCAGCCCC 833  
 RESULT 30  
 AA233574/C  
 ID AA233574 standard; cDNA; 1036 BP.  
 XX  
 AC AA233574;  
 XX  
 DT 08-DEC-1999 (first entry)  
 XX  
 DE Human breast tumour-associated EST 34.  
 XX  
 KW Expressed sequence tag; EST; human; breast; cancer; cytostatic;  
 KW medicaments; gene therapy; treatment; fat metabolism; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN DE19013835-Al.  
 XX  
 PD 23-SEP-1999.  
 XX  
 PF 20-MAR-1998; 98DE-01013835.  
 XX  
 PR 20-MAR-1998; 98DE-01013835.  
 XX  
 PA (META-) METAGEN CBS GENOMFORSCHUNG MBH.  
 XX  
 PI Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E, Rosenthal A;  
 XX  
 DR WPI; 1999-528979/45.  
 XX  
 PT Human nucleic acid sequences and protein products from normal breast  
 PT tissue, useful for breast cancer therapy.  
 XX  
 PS Claim 3; 122; 206pp; German.  
 XX  
 CC This invention describes novel human nucleic acid sequences from normal  
 CC breast tissue which have cytostatic activity. The nucleic acid sequences  
 CC can be used to produce and isolate full-length gene sequences. They can  
 CC be used to express proteins, which can be used as tools to find an  
 CC activity against breast cancer. The sequences can be used in sense or  
 CC antisense form. They are especially useful for medicaments for gene  
 CC therapy to treat breast cancer and for treating illnesses associated with  
 CC fat metabolism. AA233541-233610 represent expressed sequence tags  
 CC described in the method of the invention  
 XX  
 SQ Sequence 1036 BP; 240 A; 251 C; 283 G; 262 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.62e-36 Length: 1036  
 Score: 354.00 Matches: 66  
 Percent Similarity: 78.45% Conservative: 25  
 Best Local Similarity: 56.90% Mismatches: 19

Query Match: 55.40% Indels: 6  
 DB: 2 Gaps: 2  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x AA233574 (1-1036)  
 Qy . 8 AspSerGlyIleSerGluValArgSerAspArgAspIysPheValIlePheLeuAspVal 27  
 Db 524 GACACTGGACTCTCAGAGATCGCGCTGGAGAACAGACAGGTTCTCTGTCAACCTGGATGTG 465  
 Qy 28 LysHisPheSerProGluAspLeuThrValIysValGlnAspAspPheValGluIleHis 47  
 Db 464 AAGCACTTCTCCAGAGAACTCAAGTTAAGGTGTGGAGATGTGATTGAGGTGCAT 405  
 Qy 48 GlyIysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 67  
 Db 404 GGAACATGAAGAGCCAGGATGAACATGTTTCATCTCCAGGAGTTCACAGAAA 345  
 Qy . 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87  
 Db 344 TACCGATCCAGCTGATGTAGACCTCTCACCATTTCTATCTCCCTGTCTCATCTGATGG 285  
 Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107  
 Db 284 GTCTCTACTGTGAATGACCAAGAAACAG-----GTCTCTGGCCCTGAGCGC 237  
 Qy 108 AlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
 Db 236 ACCATTCCCATCACCCTGGAAGAGAGCTGTCTGTCAACCGCAGCCCC 189  
 RESULT 31  
 ABX38978  
 ID ABX38978 standard; cDNA; 380 BP.  
 XX  
 AC ABX38978;  
 XX  
 DT 20-FEB-2003 (first entry)  
 XX  
 DE Bovine EST associated with lactation/muscle/fat deposition #4143.  
 XX  
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
 KW muscle deposition; fat deposition; genome mapping; gene identification;  
 KW gene analysis; cattle breeding.  
 XX  
 OS Bos Taurus.  
 XX  
 PN US2002137139-A1.  
 XX  
 PD 26-SEP-2002.  
 XX  
 PF 24-SEP-2001; 2001US-00960352.  
 XX  
 PR 12-JAN-1999; 99US-0115707P.  
 PR 11-JAN-2000; 2000US-00480902.  
 XX  
 PA (BYAT/) BYATT J C.  
 PA (MATH/) MATHIALAGAN N.  
 PA (TAON/) TAO N.  
 PA (WARR/) WARREN W C.  
 XX  
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
 XX  
 DR WPI; 2003-110599/10.  
 XX  
 PT New nucleic acid associated with lactation, and muscle and fat  
 PT deposition, useful for genome mapping, gene identification and analysis,  
 PT cattle breeding, or for genetically improving cattle.  
 XX  
 PS Claim 2; SEQ ID NO 4143; 245pp; English.  
 XX  
 CC The invention relates to a purified nucleic acid molecule associated with  
 CC lactation or muscle and fat deposition (designated LMFD), derived from  
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second  
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,

CC appearing as ABX34836-ABX49947, or complements of them. Also included are  
CC ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic  
CC acid linked to a promoter and a 3' non-translated sequence that  
CC functions in the cell to cause termination of transcription and addition  
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
CC (2) determining a level or pattern of a molecule in a bovine cell or  
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any  
CC of the 15112 nucleic acid sequences or its complement or fragment) with a  
CC complementary nucleic acid molecule obtained from the bovine cell or  
CC tissue, where hybridisation between the marker nucleic acid and the  
CC complementary nucleic acid permits the detection of the molecule; and (b)  
CC detecting the level or pattern of the complementary nucleic acid, where  
CC the detection of the complementary nucleic acid is predictive of the  
CC level or pattern of the molecule. The LMPD nucleic acid is used for  
CC determining a level or pattern of a molecule in a bovine cell or tissue.  
CC It is useful for genome mapping, gene identification and analysis, cattle  
CC breeding, preparation of constructs for use in cattle gene expression, or  
CC for genetically improving cattle. The present sequence is one of the  
CC 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The  
CC present sequence was not shown in the specification but was obtained in  
CC electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?docID=20020137139  
XX  
SQ Sequence 380 BP; 90 A; 114 C; 101 G; 75 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.66e-34 Length: 380  
Score: 332.50 Matches: 65  
Percent Similarity: 75.65% Conservativeness: 22  
Best Local Similarity: 56.52% Mismatches: 23  
Query Match: 52.03% Indels: 5  
DB: 8 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x ABX38978 (1-380)

Qy 4 ArgThrValLeuAsp---SerGlyIleSerGluValArgSerAspArgAspLysPheVal 22  
Db 12 CGCACCCAGCTGGATTGCTCTGCCCTCAGAGATGCCCTGGAGAGACAGATCTCT 71  
Qy 23 IlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAsp 42  
Db 72 GTCAACCTGGACCTGACGCACTCTCTCCAGAGGAACTCAAGGCCAAGTCTGGAGAT 131  
Qy 43 PheValGluIleHisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArg 62  
Db 132 GTGATTGAGTGCATGCGCAACATCAAGAGCGCCAGGATGAACATGGTTTATCTCCCGG 191  
Qy 63 GluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSer 82  
Db 192 GAGTTCACAGAAATACCGATCCAGCTGACGTGGACCTCTCGCCATTACTTCAATCC 251  
Qy 83 LeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAla 102  
Db 252 CTGTCCTCTGATGGGACCTACTGTGATGAGCCAGCAAGAACAG-----GCC 299  
Qy 103 ThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro 117  
Db 300 TCGGCCCTGAGCGCACCATTCCTCCATACCCGTTGAAGAGAGCGC 344

RESULT 32

AAC03894

ID AAC03894 standard; cDNA; 695 BP.

XX AAC03894;

XX

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 3892.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

OS

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GBST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG03888.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1; SEQ ID NO 3892; 71pp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. An ORF has been identified within the  
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
CC derived from 30 different tissues. EST sequences usually correspond  
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
CC well suited for isolating cDNA sequences derived from the 5' ends of  
CC mRNAs and even in those cases where longer cDNA sequences have been  
CC obtained the full 5' UTR is rarely included. 5' ESTs are derived from  
CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
CC gene therapy and chromosome mapping procedures. They are used to obtain  
CC upstream regulatory sequences and to design expression and secretion  
CC vectors

XX SQ Sequence 695 BP; 183 A; 165 C; 162 G; 183 T; 0 U; 2 Other;

Alignment Scores:  
Pred. No.: 1.36e-33 Length: 695  
Score: 330.00 Matches: 62  
Percent Similarity: 77.27% Conservativeness: 23  
Best Local Similarity: 56.36% Mismatches: 19  
Query Match: 51.64% Indels: 6  
DB: 3 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x AAC03894 (1-695)

Qy 14 ValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGlu 33  
Db 269 ATGGCCCTGGAGAGACAGGTTCTCTCAACCTGGATGTGAAGCACTTCTCCACAG 328

Qy 34 AspleuThrValLysValGlnAspPheValGluIleHisGlyLysHisAsnGluArg 53  
Db 329 GAACCTCAAGTTAAGGTGTGGGAGATGTGATCAGGTGTCATGGAACATCAAGAGCGC 388

Qy 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 73  
Db 389 CAGGATGAACATGTTTCACTCCAGGAGTTCACAGAGAAATACCGGATCCAGCTGAT 448

Qy 74 ValAspGlnSerAlaLeuSerCysSerIleuSerAlaAspGlyMetLeuThrPheCysGly 93  
Db 449 GTAGACCTCTCACCATTACTTCTCTGTCATCTGATGGGTCTCTCACTGTGAATGGA 508

Qy 94 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113  
Db 509 CCAAGGAACACAG-----GTCTCTGGCCCTGAGCGCACCATTTCCATCCACCGT 556

Qy 114 GluGluLysPro-----ThrSerAlaPro 121

Db 557 GAAGAGAGAGCTGTGTGTCAACCGCACCC 586

Query Match:	60.00%	Mismatches:	11
DB:	44.29%	Indels:	0
Gaps:	8	Gaps:	0
US-10-657-740-1_COPY_51_173 (1-123) x ABX39112 (1-449)			
QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspIleValIlePheLeuAsp 26			
DB 210 ATTGACACTGGCTCTCAGAGATCCGCTCGGAGAGACAGATTCCTCTCAACCTGGAT 269			
QY 27 VallysHisPheSerProGluAspLeuThrValIleValGlnAspPheValGluIle 46			
DB 270 GTGAAGCACTTCTCCCGAGAGAACTCAAGGTCTCGGAGATGTGATTGAGGTG 329			
QY 47 HisGlyIleHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 66			
DB 330 CATGGCAACATGACAGAGGCCAGGATGACATGGTTTATCTCCGGGAGTTCACAGG 389			
QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86			
DB 390 AAATACCGGATCCGACGTGACGTGGAACCTCTCGCCATTACTCTCATCCCTGCTCTGAT 449			
RESULT 34			
ID' ABK87386			
XX ABK87386 standard; cDNA; 349 BP.			
AC ABK87386;			
XX 24-SEP-2002 (first entry)			
DT Target molecule human alpha crystallin B chain partial cDNA.			
DE Retinoic acid pathway; RA; ss; retinoid; lung cancer; Kaposi's sarcoma;			
XX breast cancer; pancreatic cancer; neuroblastoma; renal cancer;			
XX ovarian cancer; dermatitis; hyperkeratosis; eczema; Darier's disease;			
XX Reiter's disease; psoriasis; gene therapy; R2;			
XX acute promyelocytic leukaemia; APL; Perturbagen; R3; F802; F820;			
XX yeast two-hybrid assay; PAT1; kinesin light chain-related protein; human;			
XX alpha crystallin B chain.			
OS Homo sapiens.			
XX Key			
FT CDS			
FT Location/Qualifiers			
FT 1..347			
FT /*tag= a			
FT /product= "Partial alpha crystallin B chain"			
FT /partial			
FT /note= "No stop codon shown"			
FT /transl_except= (pos:1..47,aa:Ser-Val)			
XX W0200240719-A2.			
XX 23-MAY-2002.			
XX 17-NOV-2001; 2001WO-US044039.			
XX 17-NOV-2000; 2000US-0249468P.			
XX (DELT-) DELTAGEN PROTEOMICS INC.			
XX Kamb CA, Richards BT, Karpilow J;			
XX WPI; 2002-519386/55.			
XX P-PSDB; AAU99185.			
XX Polypeptide with retinoic acid pathway activity, especially of			
XX perturbagens R3, F802 and F820 for identifying a cellular target which			
XX interacts with the polypeptide and for therapeutic purposes.			
XX Example 5; Fig 16a; 131pp; English.			
XX The invention relates to an isolated polypeptide (I) with retinoic acid			
XX (RA) pathway activity, comprising a polypeptide sequence of Perturbagen			

CC (phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their  
 CC biologically active modifications, or biologically active fragments. Also  
 CC included are the polynucleotides encoding the perturbagens, a vector  
 CC comprising the polynucleotide, preparing an RA pathway related  
 CC polypeptide, a composition comprising the polypeptide, an antibody to the  
 CC polypeptide, screening (M) putative RA-related therapeutics, by exposing  
 CC a polypeptide/target interaction pair identified by a method utilising  
 CC the RA pathway polypeptide, to a number of agents, and recovering a  
 CC subpopulation of disrupting agents which competitively displace the  
 CC polypeptide from the target, where the disrupting agents are putative RA-  
 CC related therapeutics. Also include are an isolated RA pathway polypeptide  
 CC comprising PARI (a kinesin light chain-related protein) polypeptide and  
 CC its encoding polynucleotide, a gene therapy vector comprising the RA  
 CC pathway protein polypeptide or encoding or PARI and a host cell  
 CC comprising the gene therapy vector. The RA pathway polypeptide is useful  
 CC for identifying a cellular target that interacts with RA pathway-related  
 CC polypeptide, by exposing the polypeptide in vitro to putative target  
 CC molecules and identifying a polypeptide/target interaction pair, by  
 CC detecting reporter expression, where the reporter expression is  
 CC operatively linked to the formation of the interaction pair. (M) is a  
 CC yeast two-hybrid assay. The polypeptide is also useful for treating an RA  
 CC pathway-related condition e.g. lung cancer, Kaposi's sarcoma, breast  
 CC cancer, pancreatic cancer, neuroblastoma, renal cancer, ovarian cancer,  
 CC dermatitis, hyperkeratosis, eczema, Darier's disease, Reiter's disease,  
 CC psoriasis, acute promyelocytic leukaemia (APL). The present sequence is a  
 CC cDNA fragment encoding a partial target molecule isolated by the yeast  
 CC two hybrid system, alpha crystallin B chain

XX  
 SQ Sequence 349 BP; 75 A; 98 C; 87 G; 89 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3,6e-25 Length: 349  
 Score: 263.00 Matches: 44  
 Percent Similarity: 88.24% Conservative: 16  
 Best Local Similarity: 64.71% Mismatches: 8  
 Query Match: 41.16% Indels: 0  
 DB: 6 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x ABK87386 (1-349)

QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27  
 DB 132 GACACTGGACTCTCAGAGATGCGCTGGAGAGGACAGGTTCTCTGTCACCTGGATGG 191  
 QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 47  
 DB 192 AAGCACTTCTCCAGAGGAACCTCAAGTTAAGTGTTCGGAGATGCTGATGAGTGCAT 251  
 QY 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 67  
 DB 252 GGAAACACATGAAGAGCGCCAGGATGAACATGTTTCATCTCCAGGAGTTCACAGGAAA 311  
 QY 68 TyrArgLeuProSerAsnValAsp 75  
 DB 312 TACCGATCCAGCTGATGATAGAC 335

RESULT 35  
 ADM94741

ID ADM94741 standard; cDNA; 615 BP.

AC ADM94741;

DT 17-JUN-2004 (first entry)

DE Human Hsp86 (heat shock protein 27) cDNA.

XX anticancer; head; neck cancer; cytostatic; cisplatin-resistant;

KW antisense therapy; human; ss; Hsp27; heat shock protein 27.

XX Homo sapiens.

XX KR2003065206-A.

XX

PD 06-AUG-2003.

XX 31-JAN-2002; 2002KR-00005714.

XX 31-JAN-2002; 2002KR-00005714.

XX (DNAD-) DNA INC.

XX Jung DG, Kim SB, Lee DS, Park JU;

XX WPI; 2004-117343/12.

XX Antisense oligonucleotide useful in the treatment of cisplatin-resistant

XX head and neck cancer.

XX Claim 4; SEQ ID NO 2; 18pp; Korean.

XX The invention relates to a novel anticancer activity-increasing agent for  
 CC enhancing a therapeutic effect on head and neck cancer by inhibiting a  
 CC gene whose expression is increased in head and neck cancer cells. The  
 CC agent of the invention demonstrates cytostatic activities and may be  
 CC useful for enhancing a therapeutic effect on head and neck cancer by  
 CC inhibiting a gene whose expression is increased in head and neck cancer  
 CC cells, as well as for effectively treating cisplatin-resistant head and  
 CC neck cancer, possibly via antisense therapy. The current sequence is that  
 CC of the human Hsp27 (heat shock protein 27) cDNA of the invention.

XX Sequence 615 BP; 105 A; 234 C; 188 G; 88 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1,28e-24 Length: 615  
 Score: 261.50 Matches: 55  
 Percent Similarity: 65.18% Conservative: 18  
 Best Local Similarity: 49.11% Mismatches: 30  
 Query Match: 40.92% Indels: 10  
 DB: 12 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x ADN31647 (1-615)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26

DB 241 CTCACGACGGGGTCTCGGAGATCGGCACACTGCGGACCGTGGCGCGTGTCCCTGGAT 300

QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIle 46

DB 301 GTCAACCACTTCGCCCGGACGAGCTGACGGTCAAGACCAAGATGGCGTGTGGAGATC 360

QY 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 66

DB 361 ACCGCAACGACGAGGAGCGGCGGACGAGCATGCTACATCTCCGGTGTCTTCACGCGG 420

QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86

DB 421 AAATACAGCTGCCCCCGGTGTGGACCCACCAAGTTCTCTCTCTGTCTCCCTGAG 480

QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106

DB 481 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 512

QY 107 ArgAlaIleProValSerArgGluLysProThr 118

DB 513 GCTAGCCAGCAGTCCAAAGAGATCATCCATCCCACT 548

RESULT 36

ADM94741

ID ADM94741 standard; cDNA; 764 BP.

XX ADM94741;

XX 01-JUL-2004 (first entry)

XX Human heat shock protein 27 (hsp27) cDNA sequence SEQ ID NO:91.

XX

```
KW heat shock protein 27; hsp27; cytosolic; gene therapy;
KW heat shock protein 27 inhibitor; hsp27 inhibitor; cancer; human; gene;
XX ss.
XX Homo sapiens.
XX WO2004030660-A2.
XX 15-APR-2004.
XX 02-OCT-2003; 2003WO-CA001588.
XX 02-OCT-2002; 2002US-0415859P.
XX 18-APR-2003; 2003US-0463952P.
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX Gleave ME, Rocchi P, Signaevsky M;
XX WPI; 2004-316331/29.
XX New composition comprising a therapeutic agent that reduces the amount of
XX active hsp27 in hsp27 expressing cells exposed to the therapeutic agent,
XX useful in treating cancer, e.g., prostate cancer or a central nervous
XX system malignancy.
XX Disclosure; SEQ ID NO 91; 38pp; English.
XX The present invention describes a composition which comprises a
XX therapeutic agent that reduces the amount of active heat shock protein 27
XX (hsp27) in hsp27 expressing cells exposed to the therapeutic agent. The
XX composition has cytostatic activity, and can be used in gene therapy. The
XX composition is useful in treating cancer, e.g., prostate, bladder, lung,
XX breast, pancreatic, colon, skin (for example melanoma), renal or ovarian
XX cancer or a central nervous system malignancy. The present sequence
XX represents a cDNA sequence of human hsp27, which is used in the
XX exemplification of the present invention.
XX SQ Sequence 764 BP; 136 A; 283 C; 219 G; 126 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.76e-24 Length: 764
Score: 261.50 Matches: 55
Percent Similarity: 65.18% Conservative: 18
Best Local Similarity: 49.11% Mismatches: 30
Query Match: 40.92% Indels: 10
DB: 12 Gaps: 1
US-10-657-740-1_COPY_51_173 (1-123) x ADM94741 (1-764)
QY 7 LeuAspSerGlyTleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 CTGACGAGCGGGGTCTCGGAGATCCGGACACTGGCGACCGGTGGCGGTCTCGAT 325
QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValLle 46
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 GTCAACCACTTCGCCCGGACGAGGTGCGTCAAGACCAAGATGGCGGTGGAGATC 385
QY 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 66
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
386 ACCGGCAAGCAGCAGGAGCGGCGAGCATGCTCATCTCCGGTGTCTTACCGGG 445
QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
446 AATACACCGTCCCGCGGTGGAGCCCGACCCCAAGTTCTCTCCCTGTCCCTCGAG 505
QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
506 GGCACACTGACGTGGA-GGCCCC-----CATGCCCAA 537
QY 107 ArgAlaIleProValSerArgGluLysProThr 118
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
538 GCTAGCCACGAGTCCACGAGATCACCATCCAGT 573
```

## RESULT 37

```
ADI31974
ID ADI31974 standard; cDNA; 789 BP.
XX AC ADI31974;
XX DT 17-JUN-2004 (first entry)
XX DE Human cDNA #1300.
XX KW Human; gene; ss; immunological response; immunopathological condition;
KW Crohn's disease; asthma; ulcerative colitis; hyper eosinophilia;
KW Irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW acute monocytic leukaemia; antiinflammatory; antitumour; antiulcer;
KW osteopathic; antiarthritic; antirheumatic; cytostatic.
XX OS Homo sapiens.
XX PN US6607879-B1.
XX 19-AUG-2003.
XX PF 09-FEB-1998; 98US-00023655.
XX PR 09-FEB-1998; 98US-00023655.
XX (INCY-) INCYTE CORP.
XX Cocks BG, Stuart SG, Seilhamer JJ;
XX WPI; 2003-895307/82.
XX A composition comprising a plurality of cDNAs, useful for detecting
XX altered expression of genes in an immunological response or for
XX diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
XX or osteoarthritis.
XX Claim 1; SEQ ID NO 1300; 50pp; English.
XX The invention relates to a composition comprising a plurality of cDNAs
XX for detecting the altered expression of genes in an immunological
XX response. The invention also relates to a method of diagnosing or
XX monitoring the treatment of an immunopathological condition in a sample,
XX comprising obtaining nucleic acids from a sample, contacting the nucleic
XX acids of the sample with an array comprising the plurality of cDNAs under
XX conditions to form one or more hybridisation complexes, detecting the
XX hybridisation complexes and comparing the levels of the detected
XX hybridisation complexes with the level of hybridisation complexes
XX detected in a non-diseased sample, where an altered level of the detected
XX hybridisation complexes correlates with the presence of an
XX immunopathological condition. Also disclosed are an expression profile
XX comprising a microarray and a plurality of detectable complexes and a
XX method for identifying a plurality of polynucleotide probes. The cDNAs
XX are useful as hybridisable array elements in a microarray for monitoring
XX the expression of target polynucleotides. The microarray can be used in
XX the diagnosis of an immunopathology, such as Crohn's disease, asthma,
XX ulcerative colitis, hyper eosinophilia, irritable bowel syndrome,
XX osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
XX identifying agents for the treatment of the diseases. The microarray may
XX also be used in drug discovery and development, toxicological and
XX carcinogenicity studies, forensics or pharmacogenomics. The composition
XX may also be used in purification of a subpopulation of mRNAs, cDNAs or
XX genomic fragments. This sequence represents a human cDNA of the
XX invention. Note: This sequence data for this patent did not form part of
XX the printed specification but was obtained in electronic format directly
XX from USPIO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 789 BP; 140 A; 296 C; 222 G; 131 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.84e-24 Length: 789
Score: 261.50 Matches: 55
```

Percent Similarity: 65.18% Conservative: 18  
 Best Local Similarity: 49.11% Mismatches: 30  
 Query Match: 40.92% Indels: 10  
 DB: 11 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x ADI31974 (1-789)

Qy 7 LeuAspSerGlyIleSerGluValArgSerArgAspLysPheValIlePheLeuAsp 26  
 Db 293 CTCAGCAGCGGGTCTCGAGATCCGGACACTGCGGACCGCTGGCGGTGCTCCCTGGAT 352  
 Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspAspPheValGluIle 46  
 Db 353 GTCACCACTTCGCCCGGAGAGTACGTCAGACCAAGATGGCGTGGTGAGATC 412  
 Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyIleSerArgGluPheHisArg 66  
 Db 413 ACCGCAAGCAGCAGGAGCGGCGAGCAGCAGCATGCTACATCTCCCGTGTTCACGGCG 472  
 Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 473 AAATACACGCTGCCCGCGTGTGGACCCCAAGTTTCTCTCTCTCTCTCTCTCTCTGAG 532  
 Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 533 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 564  
 Qy 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
 Db 565 GCTAGCCAGCGAGTCCACGAGATCACCATCCAGT 600

RESULT 38  
 ADG10693  
 ID ADG10693 standard; cDNA; 847 BP.  
 XX AC ADG10693;  
 XX DT 26-FEB-2004 (first entry)  
 XX DE Human STAT6-activating protein-encoding cDNA, SEQ ID NO:283.  
 XX KW Human; STAT6-activating protein; drug screening; activator; inhibitor;  
 KW allergic disease; inflammation; autoimmune disease; diabetes;  
 KW hyperlipidaemia; cancer; infection; HIV infection;  
 KW human immunodeficiency; cancer; Th1 hyperfunction; antiallergic;  
 KW antiinflammatory; antidiabetic; antilipemic; antinfective; anti-HIV;  
 KW cytostatic.; gene therapy; antisense therapy; ribozyme therapy; gene; ss.  
 XX OS Homo sapiens.  
 XX PN WO200296943-A1.  
 XX PD 05-DEC-2002.  
 XX PF 22-MAY-2002; 2002MO-JP004949.  
 XX PR 25-MAY-2001; 2001JP-00157043.  
 PR 30-AUG-2001; 2001JP-00260691.  
 PR 10-OCT-2001; 2001JP-00313175.  
 XX PA (ASAH) ASAHI KASEI KOGYO KK.  
 XX PI Honda G, Mameda A, Muramatsu S, Ishizawa K;  
 XX WPI; 2003-140442/13.  
 DR P-PSDB; ADG10694.  
 XX STAT6-activating proteins and encoded genes, applicable in diagnosis of  
 PT and developing drugs to treat allergic diseases, inflammations,  
 PT autoimmune diseases, diabetes, hyperlipidemia, infections e.g. HIV, and  
 PT cancer.  
 XX Claim 4; SEQ ID NO 283; 2080bp; Japanese.

XX The invention relates to 242 human STAT6-activating proteins and cDNAs  
 CC encoding them (ADG10411-ADG10894) and to sequences with 95% or more  
 CC homology to the STAT6-activating proteins and their encoding nucleic  
 CC acids. The invention also relates to recombinant vectors and host cells  
 CC comprising a STAT6-activating protein-encoding nucleic acid; the  
 CC recombinant production of a STAT6-activating protein; an antibody  
 CC specific for a STAT6-activating protein; antisense oligonucleotides and  
 CC ribozymes targeted to nucleic acids encoding a STAT6-activating protein;  
 CC methods of screening for activators or inhibitors of STAT6-activating  
 CC proteins; drug compositions comprising a modulator of STAT6-activating  
 CC protein activity or expression; and methods of treating patients by  
 CC administration of the drug compositions. The STAT6-activating proteins,  
 CC nucleic acids encoding them, and modulators of their activity or  
 CC expression are useful in the diagnosis and treatment of allergic  
 CC diseases, inflammation, autoimmune diseases, diabetes, hyperlipidaemia,  
 CC cancer, infections (e.g., HIV), cancer and disorders associated with Th1  
 CC hyperfunction. The present sequence is related to the invention.  
 XX SQ Sequence 847 BP; 160 A; 312 C; 235 G; 140 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2.04e-24 Length: 847  
 Score: 261.50 Matches: 55  
 Percent Similarity: 65.18% Conservative: 18  
 Best Local Similarity: 49.11% Mismatches: 30  
 Query Match: 40.92% Indels: 10  
 DB: 11 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x ADG10693 (1-847)

Qy 7 LeuAspSerGlyIleSerGluValArgSerArgAspLysPheValIlePheLeuAsp 26  
 Db 348 CTCAGCAGCGGGTCTCGAGATCCGGACACTGCGGACCGCTGGCGGTGCTCCCTGGAT 407  
 Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspAspPheValGluIle 46  
 Db 408 GTCACCACTTCGCCCGGAGAGTACGCTCAGACCAAGATGGCGTGGTGAGATC 467  
 Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyIleSerArgGluPheHisArg 66  
 Db 468 ACCGCAAGCAGCAGGAGCGGCGAGCAGCATGCTACATCTCCCGTGTTCACGGCG 527  
 Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 528 AAATACACGCTGCCCGCGTGTGGACCCCAAGTTTCTCTCTCTCTCTCTCTCTGAG 587  
 Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 588 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 619  
 Qy 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
 Db 620 GCTAGCCAGCGAGTCCACGAGATCACCATCCAGT 655

RESULT 39  
 ADQ86169  
 ID ADQ86169 standard; cDNA; 847 BP.  
 XX AC ADQ86169;  
 XX DT 07-OCT-2004 (first entry)  
 XX DE Human tumour-associated antigenic target (TAT) cDNA sequence #3041.  
 XX KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
 KW cancer; cell proliferative disorder; gene; ss.  
 XX OS Homo sapiens.  
 XX PN WO2004060270-A2.  
 XX DT 22-JUL-2004.

XX PF 15-OCT-2003; 2003WO-US029126.  
 XX PR 18-OCT-2002; 2002US-0418988P.  
 XX PA (GETH ) GENENTECH INC.  
 XX PA (WUTD)/ ZHU T D.  
 XX PA (ZHOU)/ ZHOU Y.  
 XX PI Wu TD, Zhou Y;  
 XX XX WPI; 2004-534300/51.  
 XX DR New nucleic acid molecule and encoded polypeptide, for diagnosing,  
 XX PT preventing or treating cell proliferative disorders such as cancer.  
 XX PS Claim 1; SEQ ID NO 3041; 5504pp; English.  
 XX CC The present invention describes an isolated tumour-associated antigenic  
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -  
 CC (c). Also described: (1) an expression vector comprising the above  
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
 CC a process for producing a polypeptide; (4) an isolated polypeptide  
 CC comprising: (a) an amino acid sequence encoded by any of the above  
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
 CC length coding region of the above nucleotide sequences; or (c) a sequence  
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
 CC an isolated antibody that binds to the above polypeptide; (7) a process  
 CC for producing the antibody; (8) an isolated oligopeptide that binds to  
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
 CC binding organic molecule that binds to the above polypeptide; (10) a  
 CC composition of matter comprising the above (chimeric) polypeptide,  
 CC antibody, oligopeptide or TAT binding organic molecule, in combination  
 CC with a carrier; (11) an article of manufacture comprising a container and  
 CC the composition of matter contained within the container; (12) methods of  
 CC inhibiting the growth of a cell that expresses the above protein, where  
 CC the growth of the cell is at least in part dependent upon a growth  
 CC potentiating effect of the above protein; (13) a method of  
 CC therapeutically treating a mammal having a cancerous tumour comprising  
 CC cells that express the above protein; (14) a method of determining the  
 CC presence of a protein in a sample suspected of containing the protein  
 CC described above; (15) methods of diagnosing the presence of a tumour in a  
 CC mammal; (16) a method for treating or preventing a cell proliferative  
 CC disorder associated with increased expression or activity of the above  
 CC protein; and (17) a method of binding an antibody, oligopeptide or  
 CC organic molecule to a cell that expresses the protein described above.  
 CC The TAT sequences have cytostatic activities, and can be used in gene  
 CC therapy. The composition and methods are useful for diagnosing,  
 CC preventing or treating cancer. The composition is also used for preparing  
 CC a medicament for the therapeutic treatment or diagnostic detection of a  
 CC cell proliferative disorder or cancer. The present sequence represents a  
 CC human TAT cDNA sequence from the present invention.  
 XX SQ Sequence 847 BP; 160 A; 313 C; 235 G; 139 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,04e-24 Length: 847  
 Score: 261.50 Matches: 55  
 Percent Similarity: 65.18% Conservative: 18  
 Best Local Similarity: 49.11% Mismatches: 30  
 Query Match: 40.92% Indels: 10  
 DB: 13 Gaps: 1  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x ADQ86169 (1-847)  
 QY 7 LeuaspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 DB 348 CTCAGACGGGGTCTCGGATCGGCACACTCGGACCGCTGGCGGTCTGCTCGAT 407

QY 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspPheValGluIle 46  
 DB 408 GTCAACCACTTGGCCCGGAGCTGACGCTCAAGACCAAGGATGGCTGTGGAGATC 467  
 QY 47 HisGlyIleHisAsnGluArgGlnAspHisGlyIleSerArgGluPheHisArg 66  
 DB 468 ACCGCAAGCAGCAGGAGCGGACGAGCATGCTACATCTCCCGGTGTTCACGGG 527  
 QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 DB 528 AATACACGCTGCCCGCGGTGGACCCCAAGTTTCTCTCTCTCTCTCTCTCTCTGAG 587  
 QY 87 GlyMetLeuThrPheCysGlyProIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 DB 588 GGCACACTGACCGTGGG-GGCCCC-----CATGCCCAA 619  
 QY 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
 DB 620 GCTAGCCACGCGAGTCCCAACGAGATCACCATCCCACT 655  
 RESULT 40  
 ID ABQ60780  
 AC ABQ60780;  
 DT 02-AUG-2002 (first entry)  
 XX Human HSBP1 cDNA sequence SEQ ID NO:4480.  
 DE Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
 KW genetic analysis; diagnostic; antisense therapy; gene; ss.  
 OS Homo sapiens.  
 XX WO2002229086-A2.  
 FN 11-APR-2002.  
 PD 02-OCT-2001; 2001WO-US030732.  
 PF 02-OCT-2000; 2000US-0237271P.  
 PR (FARB ) BAYER CORP.  
 XX Burgess C, Astle JH, Carroll B, Catino TJ, Dwivedi P, Molino GA;  
 PI Thiagalingam A, Lewis ME;  
 XX WPI; 2002-426115/45.  
 DR P-PSDB; ABB78997.  
 XX New isolated nucleic acid that is differentially expressed in cancer  
 PT tissues useful for determining the presence of colon cancer in a cell or  
 PT tissue type, and in antisense therapy.  
 XX Claim 1; Fig 2; 796pp; English.  
 CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
 CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins  
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
 CC encoded by (I) is useful for detecting cancer in a patient sample, and  
 CC for detecting the presence or absence of a polynucleotide encoded by a  
 CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived  
 CC from (I) can be used for determining the presence of a nucleic acid which  
 CC hybridises to (I), and for determining the phenotype of cells in a sample  
 CC of cells from a patient. (I) is useful for determining the presence of  
 CC colon cancer in a cell or tissue type, for determining the presence or  
 CC state of other type of cancer, in antisense therapy, to generate  
 CC macroarrays on a solid surface, to identify a chromosome on which the  
 CC corresponding gene resides, and in tissue profiling, forensics, genetic  
 CC analysis, mapping and diagnostic applications. (I) can be used to raise  
 CC antibodies, and to screen for peptide analogues and antagonists



XX according to prognosis by determining the similarity between the level of  
SQ expression of each of five genes for which markers are listed in the  
specification, in a cell sample taken from the breast cancer patient, to  
control levels of expression for each respective five genes to obtain a  
patient similarity value. The methods are useful for classifying a breast  
cancer patient according to prognosis, kits and computer program products  
are useful for data analysis using the diagnostic, prognostic and  
statistical methods of the invention. This sequence corresponds to a  
marker used in the method of the invention.

XX SQ Sequence 865 BP; 178 A; 312 C; 235 G; 140 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.1e-24 Length: 865  
Score: 261.50 Matches: 55  
Percent Similarity: 65.18% Conservative: 18  
Best Local Similarity: 49.11% Mismatches: 30  
Query Match: 40.92% Indels: 10  
DB: 13 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x ABO60780 (1-865)

QY 7 LeuaspSerGlyIleSerGluValArgSerArgAspLysPheValIlePheLeuasp 26  
Db 348 CTCAGACGGGGTCTCGAGATCCGGACACTCGCGACCGCTGCGGTGCTTCCGTGAT 407

QY 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspPheValGluile 46  
Db 408 GTCAACCACTTCCCGCGGAGCTGACGCTCAAGACCAAGGATGGCGTGGAGATC 467

QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
Db 468 ACCGCAAGCAGGAGCGGCGAGCAGCATGCTACATCTCCGGTGTTCACGCGG 527

QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaasp 86  
Db 528 AAATACACGCTGCCCGCGGTGTGACCCACCAAGTTCTCTCCTGCTGCTGAG 587

QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
Db 588 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 619

QY 107 ArgAlaIleProValSerArgGluLysProThr 118  
Db 620 GCTAGCCAGCAGTCCACGAGATCACCATCCCACT 655

RESULT 42  
ACN38656  
ID ACN38656 standard; DNA; 865 BP.  
XX AC  
XX ACN38656;  
XX  
XX 18-NOV-2004 (first entry)  
XX  
XX Tumour-associated antigenic target (TAT) cDNA DNA304710, SEQ ID NO:2308.  
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
XX tumour; diagnosis; cell proliferative disorder; breast cancer;  
XX colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
XX central nervous system cancer; bladder cancer; pancreatic cancer;  
XX cervical cancer; melanoma; leukaemia; hybridisation probe;  
XX chromosome identification; chromosome mapping; gene mapping;  
XX gene therapy; cytostatic; gene; ss.  
XX Homo sapiens.  
XX  
XX WO2004030615-A2.  
XX  
XX 15-APR-2004.  
XX  
XX 29-SEP-2003; 2003WO-US028547.  
XX  
XX 02-OCT-2002; 2002US-0414971P.



Best Local Similarity: 49.11% Mismatches: 30  
 Query Match: 40.92% Indels: 10  
 DB: 13 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x ADP23292 (1-865)

QY 7 LeuAspSerGlyIleSerProGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 DB 348 CTCAGCAGCGGGTCTCGAGATCCGGACACTGCGGACCGCTGGCGGTGTCCCTGGAT 407

QY 27 VallysHisPheSerProGluValArgSerAspValGlnAspAspPheValGluIle 46  
 DB 408 GTCAACCACTTCGCCCGGAGCGTACGCTCAGACCAAGATGGCGTGTGGAGATC 467

QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
 DB 468 ACCGCAAGCAGCAGGAGCGGAGCAGCAGTACATCTCCCGGTGCTTACGCGG 527

QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 DB 528 AATACACGCTGCCCGCGGTGGAGCCCAAGTTCTCTCCTGTCCCTGAG 587

QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 DB 588 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 619

QY 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
 DB 620 GCTAGCCAGCGAGTCCAAACGAGATCACCATCCCACT 655

RESULT 44  
 ABN97370  
 ID ABN97370 standard; DNA; 1231 BP.  
 XX AC ABN97370;  
 XX DT 13-AUG-2002 (first entry)  
 XX DE Gene #3868 used to diagnose liver cancer.  
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
 KW metastatic liver tumour; cytostatic; expression profile; disease state;  
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
 XX OS Homo sapiens.  
 XX WO200229103-A2.  
 XX PD 11-APR-2002.  
 XX PF 02-OCT-2001; 2001WO-US030589.  
 XX PR 02-OCT-2000; 2000US-0237054P.  
 XX PA (GENE-) GENE LOGIC INC.  
 XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
 XX WPI; 2002-426119/45.  
 XX DR  
 XX PT Diagnosing and detecting the progression of liver cancer, hepatocellular  
 PT carcinoma or metastatic liver tumor in a patient, involves detecting the  
 PT level of expression of two or more genes in a liver tissue sample.  
 XX PS Claim 1; SEQ ID NO 3868; 298pp; English.  
 XX CC The invention relates to a novel method for diagnosing and detecting the  
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
 CC tumour in a patient, and differentiating metastatic liver cancer from  
 CC hepatocellular carcinoma in a patient, involving detecting the level of  
 CC expression of two or more genes represented in ABN93503-ABN97455 in a  
 CC tissue sample. The method of the invention has hepatotropic, and  
 CC cytostatic activity. The method is useful for diagnosing and detecting

CC the ptogression of liver cancer, hepatocellular carcinoma and metastatic  
 CC liver carcinoma in a patient. The method is useful for identifying  
 CC expression profiles which serve as useful diagnostic markers as well as  
 CC markers that can be used to monitor disease states, disease progression,  
 CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 1231 BP; 245 A; 396 C; 321 G; 269 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.49e-24 Length: 1231  
 Score: 261.50 Matches: 55  
 Percent Similarity: 65.18% Conservative: 18  
 Best Local Similarity: 49.11% Mismatches: 30  
 Query Match: 40.92% Indels: 10  
 DB: 6 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x ABN97370 (1-1231)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 DB 732 CTCAGCAGCGGGTCTCGAGATCCGGACACTGCGGACCGCTGGCGGTGTCCCTGGAT 791

QY 27 VallysHisPheSerProGluValArgSerAspValGlnAspAspPheValGluIle 46  
 DB 792 GTCAACCACTTCGCCCGGAGCGTACGCTCAGACCAAGATGGCGTGTGGAGATC 851

QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
 DB 852 ACCGCAAGCAGCAGGAGCGGAGCGGAGCATGGCTACATCTCCCGGTGCTTACGCGG 911

QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 DB 912 AATACACGCTGCCCGCGGTGGAGCCCAAGTTCTCTCCTGTCCCTGAG 971

QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 DB 972 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 1003

QY 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
 DB 1004 GCTAGCCAGCGAGTCCAAACGAGATCACCATCCCACT 1039

RESULT 45  
 ADD70998  
 ID ADD70998 standard; DNA; 1231 BP.  
 XX AC ADD70998;  
 XX DT 15-JAN-2004 (first entry)  
 XX DE Human heat shock 27KD protein 1 gene SEQ ID NO:2.  
 XX KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;  
 KW cytostatic; gene therapy; human; gene; ds.  
 XX OS Homo sapiens.  
 XX WO2003061564-A2.  
 XX PD 31-JUL-2003.  
 XX PF 20-DEC-2002; 2002WO-US040718.  
 XX PR 21-DEC-2001; 2001US-0341815P.  
 XX PR 31-DEC-2001; 2001US-0343185P.  
 XX PA (GENE-) GENE LOGIC INC.  
 XX PA (LGBI-) LG BIOMEDICAL INST.  
 XX PI Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;

XX DR WPI; 2003-663343/62.  
 XX PT Diagnosing liver cancer cells, useful for treating liver cancer  
 XX PT associated with chronic hepatitis or cirrhosis comprises detecting the  
 XX PT level of expression in a tissue sample of one or more genes associated  
 XX PT with cancerous liver tissues.  
 XX PS Claim 1; SEQ ID NO 2; 176pp; English.  
 XX CC The present invention describes a method for diagnosing liver cancer  
 CC cells comprising detecting the level of expression in a tissue sample of  
 CC one or more genes given in the specification (see ADD70997 to ADD71105),  
 CC where differential expression of the genes is indicative of liver cancer.  
 CC Also described: (1) detecting the progression of liver cancer in a  
 CC patient; (2) monitoring the treatment of a patient with liver cancer; (3)  
 CC treating a patient with liver cancer; (4) typing a liver disease in a  
 CC patient; (5) detecting the presence or progression of liver cancer in a  
 CC patient with chronic hepatitis or cirrhosis; (6) differentiating liver  
 CC cancer related to chronic hepatitis from liver cancer related to  
 CC cirrhosis; (7) screening for an agent capable of modulating the onset or  
 CC progression of liver cancer; (8) a composition comprising at least two  
 CC oligonucleotides comprising a sequence that specifically hybridises to  
 CC any of the genes; (9) a solid support comprising the at least two  
 CC oligonucleotides; (10) a computer system comprising a database containing  
 CC information identifying the level in liver tissue of a set of genes; (11)  
 CC a method for using the computer system to present information identifying  
 CC the expression level in tissue or cell of any of the genes; and (12) a  
 CC therapeutic agent for slowing or halting the progression of liver cancer.  
 CC The methods are useful for treating liver cancer associated with chronic  
 CC hepatitis or cirrhosis. The present sequence represents a specifically  
 CC claimed human gene sequence which is used in the exemplification of the  
 CC present invention.

XX SQ Sequence 1231 BP; 245 A; 396 C; 321 G; 269 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 3,49e-24 Length: 1231  
 Score: 261.50 Matches: 55  
 Percent Similarity: 65.18% Conservative: 18  
 Best Local Similarity: 49.11% Mismatches: 30  
 Query Match: 40.92% Indels: 10  
 DB: 10 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x ADD70998 (1-1231)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 732 CTCAGCAGCGGGTCTCGGAGATCGGCACATCGCGACCGTGGCGGTGTCTCTGGAT 791  
 QY 27 ValIysHisPheSerProGluAspLeuThrValIysValGlnAspPheValGluIle 46  
 Db 792 GTCAACCACTTCGCCCGCGAGCTGCGTCAAGACCAAGATGGTGGTGAGATC 851  
 QY 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyIleSerArgGluPheHisArg 66  
 Db 852 ACCGCGACGACGAGCGCGGACGACGACGACGACGACGACGACGACGACGACGACG 911  
 QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 912 AAATACAGCTGCCCGCGGTGGTGGACCCACCCACCAAGTTCTCTCTCTCTCTCTCTGAG 971  
 QY 87 GlyMetLeuThrPheCysGlyProIysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 972 GGCACTGACCGTGA-GGCCCC-----CATGCCCAA 1003  
 QY 107 ArgAlaIleProValSerArgGluGluIysProThr 118  
 Db 1004 GCTAGCCAGCGATCCACGAGATCACCATCCAGT 1039

RESULT 46

AAA93441

ID AAA93441 standard; DNA; 1380 BP.

XX AC AAA93441;  
 XX DT 15-SEP-2003 (revised)  
 XX DT 10-JAN-2001 (first entry)  
 XX DE GFP-HSP27 fusion gene, SEQ ID NO:169.  
 XX KW Bioreactor protein; fusion protein; recognition site;  
 KW cellular targeting sequence; cellular localisation; fluorescent protein;  
 KW protease activity detection; toxin detection; cellular stress detection;  
 KW drug discovery; cell based screening; ds.  
 OS Aequorea victoria.  
 OS Mammalia.  
 OS Chimeric.  
 XX PN WO200050872-A2.  
 XX XX 31-AUG-2000.  
 XX PF 25-FEB-2000; 2000WO-US004794.  
 XX PR 26-FEB-1999; 99US-0122152P.  
 PR 08-MAR-1999; 99US-0123399P.  
 XX 12-JUL-1999; 99US-00352171.  
 XX (CELL-) CELLOMICS INC.  
 XX Giuliano KA, Kapur R;  
 DR WPI; 2000-594086/56.  
 DR P-PSDB; AAB22936.  
 XX Automated cell-based characterization of toxin by contacting cells  
 PT containing luminescent reporter molecules with test substance and  
 PT analyzing optically.  
 XX Example 11; Page 300-302; 336pp; English.  
 CC The invention relates to systems, methods and reagents for cell-based  
 CC screening or detection of compounds which affect particular biological  
 CC functions. The methods of the invention utilise fluorescent bioreactor  
 CC molecules which, when acted on by a compound of interest, cause an  
 CC alteration in the cellular distribution of at least the fluorescent  
 CC moiety. In one embodiment, the biosensors comprise heat shock proteins  
 CC (HSPs) fused to a fluorescent protein (e.g., jellyfish green fluorescent  
 CC protein (GFP), or derivatives thereof). Such biosensors are located in  
 CC the cytoplasm, but on stress activation translocate to the nucleus. In  
 CC another embodiment bioreactor proteins can be used to detect protease  
 CC activity. Such protease bioreactor fusion proteins comprise one or more  
 CC fluorescent proteins; a recognition signal which is cleaved by the  
 CC protease; and at least one cellular localisation signal. The latter two  
 CC components may be components of a single protein which is acted upon by  
 CC the protease, or may be from heterologous sources. Due to the  
 CC localisation signal, the bioreactor protein is localised to a particular  
 CC region of the cell. Once acted on by the protease of interest, the  
 CC fluorescent protein is cleaved from the localisation sequence, and is  
 CC free to migrate to other locations within the cell. The presence of a  
 CC second localisation signal attached to the fluorescent protein enables  
 CC the fluorescent protein to be directed to a different cellular  
 CC compartment after cleavage of the protease recognition sequence. The  
 CC change in distribution of the fluorescent protein can be detected using  
 CC imaging methods with a high degree of spatial resolution. The methods and  
 CC biosensors of the invention can be used to investigate a wide range of  
 CC cellular activities and to screen compounds which modulate these  
 CC activities. Biosensors containing a recognition site for caspase, for  
 CC example, may be used for the screening of compounds which modulate  
 CC apoptosis, while biosensors containing other protease recognition sites  
 CC may be used for the detection of proteolytic toxins (such as anthrax  
 CC lethal factor). The method provides improved target validation and  
 CC candidate compound optimisation by combining many cell screening formats  
 CC with fluorescence-based molecular reagents and computer-based feature



AC ABK63074;  
XX  
XX 18-JUN-2002 (first entry)  
XX  
XX Rat sequence differentially expressed in response to a hepatotoxin #981.  
XX  
XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;  
XX  
XX differential expression; centrilobular necrosis; steatosis.  
XX  
XX Rattus norvegicus.  
XX  
XX  
XX WO200210453-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 30-JUL-2001; 2001WO-US023872.  
XX  
XX 31-JUL-2000; 2000US-0222040P.  
XX  
XX 02-NOV-2000; 2000US-0244880P.  
XX  
XX 11-MAY-2001; 2001US-0290029P.  
XX  
XX 15-MAY-2001; 2001US-0290645P.  
XX  
XX 22-MAY-2001; 2001US-0292336P.  
XX  
XX 06-JUN-2001; 2001US-0292579P.  
XX  
XX 13-JUN-2001; 2001US-0297457P.  
XX  
XX 19-JUN-2001; 2001US-0298884P.  
XX  
XX 09-JUL-2001; 2001US-0303459P.  
XX  
XX (GENE-) GENE LOGIC INC.  
XX  
XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;  
XX  
XX WPI; 2002-241625/29.  
XX  
XX  
XX Predicting toxic effects of compounds or the progression of these toxic  
XX  
XX effects by determining the changes in gene expression in tissues or cells  
XX  
XX exposed to the toxin and comparing these to gene expression in unexposed  
XX  
XX tissues or cells.  
XX  
XX Claim 1; SEQ ID NO 981; 239pp; English.  
XX  
XX The invention relates to methods for predicting toxic effects of  
XX  
XX compounds or the progression of these toxic effects by determining the  
XX  
XX global changes in gene expression in tissues or cells exposed to the  
XX  
XX toxin and comparing these to gene expression in unexposed tissues or  
XX  
XX cells. Also included are methods of predicting at least one toxic effect  
XX  
XX of a compound or progression of a toxic effect, preferably the  
XX  
XX hepatotoxicity of a compound, comprising detecting the level of  
XX  
XX expression in a tissue or cell sample exposed to the compound of two or  
XX  
XX more genes listed in the specification, where differential expression of  
XX  
XX the genes is indicative of at least one toxic effect or progression. The  
XX  
XX method can also be used to identify an agent which modulates the toxic  
XX  
XX response and predict cellular pathways that a compound modulates in a  
XX  
XX cell. The methods utilise a set of at least two probes (on a solid  
XX  
XX support in kit form), where each of the probes comprises a sequence that  
XX  
XX specifically hybridises to a gene listed in the specification, a computer  
XX  
XX system comprising a database containing information identifying the  
XX  
XX expression level in a tissue or cell sample exposed to a hepatotoxin of a  
XX  
XX set of genes comprising at least two genes listed in the specification,  
XX  
XX and a user interface to view the information used to present information,  
XX  
XX identifying the expression level in a tissue or cell of at least one gene  
XX  
XX listed in the specification. The method is useful for elucidating global  
XX  
XX changes in gene expression and for identifying toxicity markers in  
XX  
XX tissues or cell exposed to a known toxin. The genes may be used as  
XX  
XX toxicity markers in drug screening and toxicity assays. The genes and  
XX  
XX gene expression information may be used as diagnostic markers for the  
XX  
XX prediction or identification of the physiological state of tissue or cell  
XX  
XX sample that has been exposed to a compound or agent. Hepatotoxicity is  
XX  
XX characterised by centrilobular necrosis and steatosis. The present  
XX  
XX sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
XX  
XX which is differentially expressed in response to a hepatotoxic agent  
XX  
XX  
XX Sequence 604 BP; 124 A; 156 C; 199 G; 125 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1-46e-24 Length: 604  
Score: 261.00 Matches: 57  
Percent Similarity: 65.18% Conservative: 16  
Best Local Similarity: 50.89% Mismatches: 33  
Query Match: 40.85% Indels: 6  
DB: 6 Gaps: 2  
US-10-657-740-1\_COPY\_51\_173 (1-123) x ABK63074 (1-604)  
QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
Db 492 CTCAGTAGCGGTCTCTCAGAGATCCGACAGACGGCGGATCGCTGGCGGTGCTCCCTGGAC 433  
QY 27 VallysHisPheSerProGluAspLeuThrValIysValGlnAspAspPheValGluIle 46  
Db 432 GTCAACCACTTGGCTCTCAGAGAGCTCACATTAAAGAACCAAGGAGCGGTGGTGGAGATC 373  
QY 47 HisGlyIysHisAsnGluArgGlnAspHisGlyIleSerArgGluPheHisArg 66  
Db 372 ACTGGCAAGCAGCAAGAAAGGAGGATGAAATGCTGCTACATCTCTCGGTGCTTCCCGGG 313  
QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 312 AAATACAGCTCCCTCCAGGTGGAGCCCACTTGGTGTCTCTCTCCCTGTCCTCCCTGAG 253  
QY 87 GlyMetLeuThrPheCysGly-----ProLysIleGlnThrGlyLeuAspAlaThrHis 104  
Db 252 GGCACTCAGCGTGAAGGCTCCGCTCCCAAGACAGTCA-----CAATCA 205  
QY 105 AlaGluArgAlaIleProValSerArgGluGluLys 116  
Db 204 GCGGAGATCACTTCCGTCACCTTTCGAGGCGCGT 169  
RESULT 49  
ADP72259/C  
ID ADP72259 standard; DNA; 604 BP.  
XX AC ADP72259;  
XX DT 26-AUG-2004 (first entry)  
XX DE Renal toxin progression gene marker #848.  
XX KW ds; toxic effect; gene expression profile; kidney tissue;  
XX KW differential gene expression; toxicity progression; toxicity marker;  
XX KW drug screening; toxicity assay; kidney pathology; nephritis;  
XX KW kidney necrosis; glomerular injury; tubular injury;  
XX KW focal segmental glomerulosclerosis.  
XX OS Rattus norvegicus.  
XX PN WO2004048598-A2.  
XX XX 10-JUN-2004.  
XX PF 24-NOV-2003; 2003WO-US037556.  
XX PR 22-NOV-2002; 2002US-00301856.  
XX PA (GENE-) GENE LOGIC INC.  
XX PI Mendrick D, Porter MW, Johnson KR, Castle A, Higgs B;  
XX PI Elashoff M;  
XX XX WPI; 2004-460771/43.  
XX PT Predicting (the progression of) a toxic effect of a compound, for  
XX PT monitoring the progression of renal disease states, comprises preparing a  
XX PT gene expression profile of a kidney tissue or cell sample exposed to the  
XX PT compound.  
XX PS Claim 11; SEQ ID NO 848; 266pp; English.

XX The invention relates to a method of predicting (the progression of) a  
 CC toxic effect of a compound by preparing a gene expression profile of a  
 CC kidney tissue or cell sample exposed to the compound and comparing the  
 CC gene expression profile to a database, or detecting the level of gene(s)  
 CC expression in a tissue or cell sample exposed to the compound, where  
 CC differential gene expression compared to a control indicates a toxic  
 CC effect (toxicity progression). The method is useful for predicting (the  
 CC progression of) at least one toxic effect of a compound. The genes are  
 CC useful as toxicity markers in drug screening and toxicity assays. The  
 CC methods are useful for predicting the likelihood that a compound or test  
 CC agent will induce various specific kidney pathologies, such as nephritis,  
 CC kidney necrosis, glomerular and tubular injury, or focal segmental  
 CC glomerulosclerosis. The methods are useful for determining the similarity  
 CC of a toxic response to one or more individual compounds and for  
 CC predicting or elucidating the potential cellular pathways influenced,  
 CC induced or modulated by the compound or test agent. The kit is useful for  
 CC predicting or modelling the toxic response of a test compound, for  
 CC monitoring the progression of renal disease states, for identifying genes  
 CC that show promise as new drug targets and for screening known and newly  
 CC designed drugs. This sequence corresponds to a gene marker used in the  
 CC method of the invention. (Note: The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences).

XX Sequence 604 BP; 124 A; 156 C; 199 G; 125 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.46e-24 Length: 604  
 Score: 261.00 Matches: 57  
 Percent Similarity: 65.18% Conservative: 16  
 Best Local Similarity: 50.88% Mismatches: 33  
 Query Match: 40.85% Indels: 6  
 DB: 12 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x ADP72259 (1-604)

Qy 7 LeuaspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuasp 26  
 Db 492 CTCAGTAGCGGTCTCAGAGATCCGACGCGCGATCGTGGCGGTGTCCCTGGAC 433  
 Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspPheValGlu 46  
 Db 432 GTCAACCACTTCGCTCCTGAGAGTCAAGTAAAGACCAAGAGAGCGGTGTGGAGATC 373  
 Qy 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
 Db 372 ACTGCACAGCAGAGAAAGGAGGATGACATGCTACATCTCTCGGTGCTTCAACCGG 313  
 Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 86  
 Db 312 AAATACACGCTCCCTCCAGGTGTGACCCACCTTGGTGTCTCTTCCCTGTCCCTGAG 253  
 Qy 87 GlyMetLeuThrPheCysGly-----ProlsIleGlnThrGlyLeuAspAlaThrHis 104  
 Db 252 GGCACATCAGCGGTGAGGCTCCGCTGCCAAAGCAGTCACA-----CAATCA 205  
 Qy 105 AlaGluArgAlaIleProValSerArgGluGluLys 116  
 Db 204 GCGGAGATCACCATTCCGGTCACTTTCGAGGCCCTG 169

RESULT 50

ABV94670

ID ABV94670 standard; cDNA; 599 BP.

XX ABV94670;

AC ABV94670;

XX 14-JAN-2003 (first entry)

DT Human pancreatic cancer expressed cDNA SEQ ID NO 43.

DE Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;

XX cytostatic; tumour; gene; ss.

KW

XX OS.

XX Homo sapiens.

XX WO200260317-A2.

XX PD 08-AUG-2002.

XX PF 30-JAN-2002; 2002WO-US002781.

XX PR 30-JAN-2001; 2001US-0265305P.

XX PR 31-JAN-2001; 2001US-0265682P.

XX PR 09-FEB-2001; 2001US-0267568P.

XX PR 21-MAR-2001; 2001US-0278651P.

XX PR 28-APR-2001; 2001US-0287112P.

XX PR 16-MAY-2001; 2001US-0291631P.

XX PR 12-JUL-2001; 2001US-0305484P.

XX PR 20-AUG-2001; 2001US-0313999P.

XX PR 27-NOV-2001; 2001US-0333626P.

XX PA (CORI-) CORIAX CORP.

XX PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;

XX DR WPI; 2002-627435/67.

XX PT New isolated polynucleotide and pancreatic tumor polypeptides, useful for

XX PT diagnosing, preventing and/or treating cancer, particularly pancreatic

XX PT cancer.

XX PS Claim 1; SEQ ID NO 43; 300pp + Sequence Listing; English.

XX CC The invention relates to an isolated polynucleotide (I) comprising: (a)

XX CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)

XX CC complements of (a); (c) sequences consisting of at least 20 contiguous

XX CC residues of (a); (d) sequences that hybridize to (a); under moderately

XX CC stringent conditions; (e) sequences having at least 75% or 90% identity

XX CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP88596-

XX CC ABP8637) encoded by (I) and oligonucleotide can be used to detect cancer

XX CC in a patient and compositions comprising polypeptides, polynucleotides,

XX CC antibodies, fusion proteins, T cell populations and antigen presenting

XX CC cells expressing the polypeptide are useful in treating pancreatic cancer

XX CC and stimulating an immune response. The polynucleotides can be used as

XX CC probes or primers for nucleic acid hybridisation, in the design and

XX CC preparation of ribozyme molecules for inhibiting expression of the tumour

XX CC polypeptides and proteins in the tumour cells, in vaccines and for gene

XX CC therapy. Note: The sequence data for this patent did not form part of the

XX CC printed specification, but was obtained in electronic format directly

XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 599 BP; 99 A; 228 C; 187 G; 85 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.95e-24 Length: 599

Score: 260.00 Matches: 55

Percent Similarity: 65.18% Conservative: 18

Best Local Similarity: 49.11% Mismatches: 30

Query Match: 40.69% Indels: 10

DB: 6 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x ABV94670 (1-599)

Qy 7 LeuaspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuasp 26

Db 254 CTCAGCAGCGGTCTCGAGATCCGACACACTCGGACCGTGGCGGTGTCCCTGGAT 313

Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspPheValGlu 46

Db 314 GTCAACCACTTCGCCCCCGACGAGCTGACGGTCAAGACCAAGAGTGGTGTGGAGATC 373

Qy 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66

Db 374 ACCGCAAGCAGCAGGAGCGGACGAGCATGCTACATCTCTCCCGGTGCTTACCGCG 433



Search completed: May 30, 2005, 05:14:51  
Job time : 382.402 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 30, 2005, 04:24:56 ; Search time 120.922 Seconds  
(without alignments)  
1664.391 Million cell updates/sec

Title: US-10-657-740-1\_COPY\_51\_173

Perfect score: 639

Sequence: 1 SLFRTVLDGSEIVSRDRK.....HAERAIPVSRKPTSPSS 123

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0			0.5
Ygapop 6.0			7.0
Delop 6.0			7.0

Searched: 1202784 seqs, 81813859 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO spool/US10657740/runat 27052005 165254 3343/app query.fasta\_1.590  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=150 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10657740 @CEN 1 1 219 @runat 27052005 165254 3343 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	639	100.0	1114	4	US-09-949-016-5220 Sequence 5220, Ap
2	367	57.4	7775	4	US-09-949-016-16962 Sequence 16962, A
3	354	55.4	691	4	US-09-949-016-2545 Sequence 2545, Ap
4	354	55.4	856	4	US-09-513-999C-3891 Sequence 3891, Ap
5	354	55.4	893	4	US-09-513-999C-14942 Sequence 14942, A
6	354	55.4	911	4	US-09-513-999C-14938 Sequence 14938, A
7	354	55.4	913	4	US-09-513-999C-14937 Sequence 14937, A
8	354	55.4	927	4	US-09-513-999C-14936 Sequence 14936, A
9	354	55.4	942	4	US-09-513-999C-3890 Sequence 3890, Ap
10	330	51.6	695	4	US-09-513-999C-3892 Sequence 3892, Ap
11	309	48.4	826	4	US-09-621-976-3172 Sequence 3172, Ap
12	261.5	40.9	789	4	US-09-023-655-1300 Sequence 1300, Ap

13	261.5	40.9	845	4	US-09-949-016-4992 Sequence 4992, Ap
14	261.5	40.9	1380	3	US-09-513-783A-169 Sequence 169, Ap
15	260	40.7	1379	3	US-09-553-498-5 Sequence 5, Appli
16	260	40.7	1379	3	US-09-618-869-5 Sequence 5, Appli
17	188	29.4	7125	4	US-09-949-016-14287 Sequence 14287, A
18	182.5	28.6	393	4	US-09-513-999C-10503 Sequence 10503, A
19	159	24.9	1627	2	US-08-900-407-2 Sequence 2, Appli
20	158	24.7	601	4	US-09-949-016-91317 Sequence 91317, A
21	158	24.7	763	4	US-09-949-016-2623 Sequence 2623, Ap
22	158	24.7	4763	4	US-09-949-016-14365 Sequence 14365, A
23	149.5	23.4	333	4	US-09-513-999C-10502 Sequence 10502, A
24	142	22.2	601	4	US-09-949-016-89647 Sequence 89647, A
25	123.5	19.3	599	4	US-09-270-767-563 Sequence 563, App
26	123.5	19.3	599	4	US-09-270-767-15845 Sequence 15845, A
27	119	18.6	5688	4	US-09-949-016-16734 Sequence 16734, A
28	118	18.5	695	4	US-09-669-751-69 Sequence 69, Appli
29	118	18.5	601	4	US-09-949-016-89648 Sequence 89648, A
30	115	18.0	12214	4	US-09-949-016-17284 Sequence 17284, A
31	110.5	17.3	719	4	US-09-620-312D-669 Sequence 669, App
32	110	17.2	9615	4	US-09-949-016-17282 Sequence 17282, A
33	107	16.7	7125	4	US-09-949-016-14287 Sequence 14287, A
34	97	15.2	601	4	US-09-949-016-175645 Sequence 175645, A
35	90	14.1	420	4	US-09-107-532A-670 Sequence 670, App
36	88.5	13.8	652	1	US-07-748-762-2 Sequence 2, Appli
37	88	13.8	168971	4	US-09-949-016-13807 Sequence 13807, A
38	86	13.5	489	4	US-09-902-540-6436 Sequence 6436, Ap
39	86	13.5	3626	4	US-09-902-540-444 Sequence 444, App
40	85	13.3	576	4	US-09-489-039A-3463 Sequence 3463, Ap
41	79.5	12.4	1350	4	US-09-583-110-2262 Sequence 2262, Ap
42	79.5	12.4	1353	4	US-09-107-433-2322 Sequence 2322, Ap
43	77.5	12.1	49487	4	US-09-949-016-11770 Sequence 11770, A
44	77.5	12.1	57751	4	US-09-949-016-13631 Sequence 13631, A
45	75.5	11.8	966	4	US-09-969-532-23 Sequence 23, Appli
46	75.5	11.8	999	4	US-09-969-532-21 Sequence 21, Appli
47	75.5	11.8	1008	4	US-09-969-532-19 Sequence 19, Appli
48	75.5	11.8	1041	4	US-09-969-532-17 Sequence 17, Appli
49	75.5	11.8	1659	4	US-09-969-532-7 Sequence 7, Appli
50	75.5	11.8	1692	4	US-09-969-532-5 Sequence 5, Appli
51	75.5	11.8	1701	4	US-09-969-532-3 Sequence 3, Appli
52	75.5	11.8	1734	4	US-09-969-532-1 Sequence 1, Appli
53	75.5	11.8	1968	4	US-09-969-532-31 Sequence 31, Appli
54	75.5	11.8	2001	4	US-09-969-532-29 Sequence 29, Appli
55	75.5	11.8	2010	4	US-09-969-532-27 Sequence 27, Appli
56	75.5	11.8	2043	4	US-09-969-532-25 Sequence 25, Appli
57	75.5	11.8	2661	4	US-09-969-532-15 Sequence 15, Appli
58	75.5	11.8	2694	4	US-09-969-532-13 Sequence 13, Appli
59	75.5	11.8	2703	4	US-09-969-532-11 Sequence 11, Appli
60	75.5	11.8	2736	4	US-09-969-532-9 Sequence 9, Appli
61	75.5	11.8	3411	4	US-09-969-532-33 Sequence 33, Appli
62	72.5	11.3	3402	1	US-08-480-917-1 Sequence 1, Appli
63	72.5	11.3	3402	3	US-09-138-736-1 Sequence 1, Appli
64	72.5	11.3	3402	3	US-08-988-242-1 Sequence 1, Appli
65	72.5	11.3	139936	4	US-09-949-016-11782 Sequence 11782, A
66	72.5	11.3	139952	4	US-09-949-016-13280 Sequence 13280, A
67	71.5	11.2	465	4	US-09-252-991A-12236 Sequence 12236, A
68	71.5	11.2	495	4	US-09-252-991A-12170 Sequence 12170, A
69	71.5	11.2	522	4	US-09-252-991A-12373 Sequence 12373, A
70	71.5	11.2	1354	3	US-08-456-2008-4 Sequence 4, Appli
71	71.5	11.2	3563	3	US-09-041-886-20 Sequence 20, Appli
72	71.5	11.2	3596	3	US-08-779-801-5 Sequence 5, Appli
73	71.5	11.2	3596	3	US-09-298-441-5 Sequence 5, Appli
74	71.5	11.2	3632	2	US-08-779-801-4 Sequence 4, Appli
75	71.5	11.2	3632	2	US-08-779-801-4 Sequence 4, Appli
76	71.5	11.2	3632	3	US-09-298-441-3 Sequence 3, Appli
77	71.5	11.2	3632	3	US-09-298-441-3 Sequence 3, Appli
78	71.5	11.2	7791	2	US-08-149-097D-23 Sequence 23, Appli
79	71.5	11.2	7791	3	US-08-949-386-23 Sequence 23, Appli
80	71.5	11.2	7791	3	US-08-450-562-23 Sequence 23, Appli
81	71.5	11.2	7791	3	US-08-984-709A-23 Sequence 23, Appli
82	71.5	11.2	7791	3	US-08-450-272-23 Sequence 23, Appli
83	71.5	11.2	7791	3	US-08-450-272-23 Sequence 23, Appli
84	71.5	11.2	7808	2	US-08-149-097D-22 Sequence 22, Appli
85	71.5	11.2	7808	3	US-08-949-386-22 Sequence 22, Appli

```
c 86 71.5 11.2 7808 3 US-08-450-562-22
c 87 71.5 11.2 7808 3 US-08-984-709A-22
c 88 71.5 11.2 7808 3 US-08-450-273-22
c 89 71.5 11.2 7808 4 US-08-450-273-22
c 90 71.5 11.2 7815 4 US-09-949-016-3629
c 91 71.5 11.2 7815 4 US-09-949-016-3630
c 92 71 11.1 1278 4 US-09-902-540-8457
c 93 71 11.1 8590 4 US-09-902-540-889
c 94 71 11.1 1830121 4 US-09-557-884-1
c 95 71 11.1 1830121 4 US-09-643-990A-1
c 96 71 11.1 4403765 3 US-09-103-840A-2
c 97 71 11.1 4411529 3 US-09-103-840A-1
c 98 70.5 11.0 585 4 US-09-248-796A-2425
c 99 70 11.0 601 4 US-09-949-016-142871
c 100 70 11.0 601 4 US-09-949-016-182120
c 101 70 11.0 265038 4 US-09-949-016-15779
c 102 69.5 10.9 42931 4 US-08-311-731A-129
c 103 69 10.8 2379 4 US-09-543-681A-2445
c 104 69 10.8 421491 4 US-09-949-016-12805
c 105 69 10.8 421494 4 US-09-949-016-14060
c 106 68.5 10.7 327 4 US-09-134-000C-1375
c 107 68.5 10.7 576 1 US-08-086-428B-34
c 108 68.5 10.7 576 2 US-08-468-570-34
c 109 68.5 10.7 576 2 US-08-290-665A-34
c 110 68.5 10.7 576 4 US-08-466-601A-34
c 111 68.5 10.7 576 5 PC1-US95-10398-34
c 112 68.5 10.7 609 4 US-09-248-796A-1874
c 113 68 10.6 531 4 US-09-252-991A-9454
c 114 68 10.6 804 4 US-09-902-540-8110
c 115 68 10.6 834 4 US-09-252-991A-9477
c 116 68 10.6 1425 4 US-09-252-991A-9353
c 117 68 10.6 1576 4 US-09-949-016-1340
c 118 68 10.6 2214 4 US-09-902-540-4124
c 119 68 10.6 2522 3 US-09-058-389A-1
c 120 68 10.6 2522 3 US-09-611-781-1
c 121 68 10.6 6890 4 US-09-902-540-828
c 122 68 10.6 23091 4 US-09-902-540-1204
c 123 68 10.6 1664976 4 US-08-916-421B-1
c 124 68 10.6 1664976 4 US-09-692-570-1
c 125 67.5 10.6 819 4 US-09-107-532A-1887
c 126 67.5 10.6 1185 4 US-09-248-796A-3762
c 127 67.5 10.6 5517 3 US-09-120-663-1
c 128 67.5 10.6 5517 4 US-09-431-614-7
c 129 67.5 10.6 6789 4 US-09-269-446D-43
c 130 67.5 10.6 9556 4 US-09-902-540-929
c 131 67.5 10.6 1194 4 US-09-252-991A-13276
c 132 67 10.5 1305 4 US-09-252-991A-12685
c 133 67 10.5 1545 4 US-09-252-991A-11637
c 134 67 10.5 5410 3 US-09-221-017B-70
c 135 67 10.5 20757 4 US-09-902-540-1189
c 136 67 10.5 145287 4 US-09-949-016-13530
c 137 67 10.5 145287 4 US-09-949-016-13531
c 138 66.5 10.4 1337 4 US-09-690-454-19
c 139 66.5 10.4 1930 1 US-07-982-112-1
c 140 66.5 10.4 54484 4 US-09-902-540-1272
c 141 66.5 10.4 1242 3 US-09-372-448A-3
c 142 65.5 10.3 876 4 US-09-252-991A-12184
c 143 65.5 10.3 882 4 US-09-328-352-710
c 144 65.5 10.3 1392 4 US-09-252-991A-12249
c 145 65.5 10.3 1407 4 US-09-252-991A-12361
c 146 65.5 10.3 3546 3 US-09-177-431-7
c 147 65.5 10.3 42246 4 US-09-949-016-17008
c 148 65.5 10.3 59123 4 US-09-949-016-12177
c 149 65.5 10.3 70559 4 US-09-409-800B-1
c 150 65.5 10.3 70559 4
```

## ALIGNMENTS

## RESULT 1

US-09-949-016-5220

; Sequence 5220, Application US/09949016

; Patent No. 6812339

## GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5220  
; LENGTH: 1114  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-5220

## Alignment Scores:

Pred. No.: 111e-80 Length: 1114  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-949-016-5220 (1-1114)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20

Db 220 TCCCTCTCCGACCCTGCTGGACATCCGGCATCTCTGAGTTGATCCGACCGGACAG 279

QY 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValLysValGln 40

Db 280 TTGCGTCTCTCCGATGTGAAGCATTCTCCCGAGGACCTCACCGTGAAGGTGCAG 339

QY 41 AspAspPheValGluIleHisGlyIleHisAsnGluValArgGlnAspAspHisGlyIle 60

Db 340 GAGCACTTTGTGGATCCACGGAAGCACACGAGCGCGGACGACCGCTACAT 399

QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80

Db 400 TCCCTGAGTTCCACGCGCGCTACCGCTCCAGTCCAGTCCAGTCCGCTCTCT 459

QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100

Db 460 TGCTCCCTGCTGCGCATGGCATGCTGACCTTCTGTGGCCCCCAAGATCCAGATCGGCTG 519

QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120

Db 520 GATGCCACCCAGCGGAGCGGATCCCGTGTCCGGGAGGAGAGGCCACCTCGGCT 579

QY 121 ProSerSer 123

Db 580 CCTCGTCC 588

## RESULT 2

US-09-949-016-16962

; Sequence 16962, Application US/09949016

; Patent No. 6812339

## GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16962
; LENGTH: 775
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16962

Alignment Scores:
Pred. No.: 9,17e-41 Length: 775
Score: 367.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.43% Indels: 0
DB: 4 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-09-949-016-16962 (1-7775)
QY 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 73
Db 5039 CAGGACGACCGGTACATTTCCCGTGAGTTCCACCGCGCTACCGCTGCGTCCAC 5098
QY 74 ValAspGlnSerAlaLeuSerCysSerIleuSerAlaAspGlyMetLeuThrPheCysGly 93
Db 5099 GTGGACCAAGTCGGCCCTCTTGTGCTCCCTGTGCGGATGGCATGCTGACCTTCTGTGCG 5158
QY 94 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluAtrGAlaIleProValSerArg 113
Db 5159 CCCAGATCCAGACGTGGCTGTGATGCCACCCACGCGAGCGCATCCCGGTGCGCG 5218
QY 114 GluGluLysProThrSerAlaProSerSer 123
Db 5219 GAGGAGAAGCCACCTCGGCTCCCTCGTCC 5248

RESULT 3
US-09-949-016-2545
; Sequence 2545, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2545
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2545

Alignment Scores:
Pred. No.: 1.62e-40 Length: 691
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservative: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 4 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x US-09-949-016-2545 (1-691)
QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 8 GACTGGACTCTCAGATGCGCTGGAGAGGACAGGTTCTGTCAACCTGGATGG 268
QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIleHis 47
Db 269 AAGCACTTCTCCCGACAGGAACTCAAGTTAAGGTGTTGGGAGATGTGATTGAGGTGCAT 328
QY 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgArg 67
Db 329 GGAACATGAAAGAGCGCCAGGATGAACATGTTTTCATCTCCAGGAGATTCCACAGAAA 388
QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerIleuSerAlaAspGly 87
Db 389 TACCGATCCAGCTGATGATAGACCCCTCTCACCATTACTTCACTCCTGTCACTGTATGGG 448
QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 449 GTCTCTCACTGTGAATGGACCAAGAAACAG-----GTCTCTGGCCCTGAGCGC 496
QY 108 AlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121
Db 497 ACCATTCCATCACCCTGTGAAGAGAGCTGTGTGTCAGCGCAGCCCCC 544

RESULT 4
US-09-513-999C-3891
; Sequence 3891, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 3891
; LENGTH: 856
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..753
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 51
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 117
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 118
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 159
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 160
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 182
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
```

```

; LOCATION: 183
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 16
; OTHER INFORMATION: Xaa=Asp or Glu
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 38
; OTHER INFORMATION: Xaa=His or Gln
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 39
; OTHER INFORMATION: Xaa=Ala or Ser
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 53
; OTHER INFORMATION: Xaa=Glu or Gln
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 60
; OTHER INFORMATION: Xaa=Glu or Gly
US-09-513-999C-3891

```

```

Alignment Scores:
Pred. No.: 2,24e-40 Length: 856
Score: 354.00 Matches: 65
Percent Similarity: 78.45% Conservative: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 4 Gaps: 2

```

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-513-999C-3891 (1-856)

```

QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
DB 412 GACACTGGACTCTCAGAGATCGCCTGGAGAGACAGGTTCTCTGTCACACCTGGATGTG 471
QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLeuHis 47
DB 472 AGGCACCTTCTCCCGAGAGAACTCAAGTTAAGGTGTGGAGATGTGATTTGAGTGCAT 531
QY 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgArg 67
DB 532 GGAATAACATGAAGAGCGCCAGCATGAACATGTTTCATCTCCAGGAGTTCACAGGAAA 591
QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
DB 592 TACCGATCCCGATGATGATAGACCTCTCACCATTACTTCCATCTCTGTCATCTGATGG 651
QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
DB 652 GTCTCTCACTGTGAATGACCAAGAAACAG-----GTCTCTGGCCCTGAGCGC 699
QY 108 AlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121
DB 700 ACCATTCCCATCACCCTGAGAGAGAGCGCTGTGTCTCAGCGAGCCGCC 747

```

#### RESULT 5

```

US-09-513-999C-14942
; Sequence 14942, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26

```

```

; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14942
; LENGTH: 893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-14942

```

```

Alignment Scores:
Pred. No.: 2,39e-40 Length: 893
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservative: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 4 Gaps: 2

```

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-513-999C-14942 (1-893)

```

QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
DB 449 GACACTGGACTCTCAGAGATCGCCTGGAGAGACAGGTTCTCTGTCACACCTGGATGTG 508
QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLeuHis 47
DB 509 AAGCACCTTCTCCCGAGAGAACTCAAGTTAAGGTGTGGAGATGTGATTTGAGTGCAT 568
QY 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgArg 67
DB 569 GGAATAACATGAAGAGCGCCAGCATGAACATGTTTCATCTCAGGAGTTCACAGGAAA 628
QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
DB 629 TACCGATCCCGATGATGATAGACCTCTCACCATTACTTCCATCTCTGTCATCTGATGG 688
QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
DB 689 GTCTCTCACTGTGAATGACCAAGAAACAG-----GTCTCTGGCCCTGAGCGC 736
QY 108 AlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121
DB 737 ACCATTCCCATCACCCTGAGAGAGAGCGCTGTGTCTCAGCGAGCCGCC 784

```

#### RESULT 6

```

US-09-513-999C-14938
; Sequence 14938, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14938
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-14938

```

```

Alignment Scores:
Pred. No.: 2,47e-40 Length: 911
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservative: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 4 Gaps: 2

```



Db 543 AAGCACTTCTCCAGAGGAACCAAGTTAAGGTGGAGATGGATTGAGTGCAT 602  
QY 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 67  
Db 603 GGAACAATCAAGAGCCCGAGATGAACATGGTTTCACTCCAGGGAGTTCCACAGAAA 662  
QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87  
Db 663 TACCGATCCAGCTGATGTAGACCTCTCACCATTAATCTATCTCTGTCATCTGATGG 722  
QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107  
Db 723 GTCTCTACTGTGAATGACCAAGAGAAACAG-----GTCTCTGCGCTGAGCGC 770  
QY 108 AlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
Db 771 ACCATTCCCATCACCCTGGAGAGAACCTTGCTGTCCCGAGCCCCC 818  
RESULT 9  
US-09-513-999C-3890  
; Sequence 3890, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 3890  
; LENGTH: 942  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 315..839  
US-09-513-999C-3890  
Alignment Scores:  
Pred. No.: 2,59e-40 Length: 942  
Score: 354.00 Matches: 66  
Percent Similarity: 78.45% Conservative: 25  
Best Local Similarity: 56.90% Mismatches: 19  
Query Match: 55.40% Indels: 6  
DB: 4 Gaps: 2  
US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-513-999C-3890 (1-942)  
QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27  
Db 498 GACACTGACTCTCAGATGCGCTGGAGAGGACAGGTTCTCTGTCAACTGGATGTG 557  
QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 47  
Db 558 AAGCACTTCTCCAGAGGAACCTCAAGTTAAGGTGGAGATGTGATGTGATGCAT 617  
QY 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 67  
Db 618 GGAACAATCAAGAGCCCGAGATGAACATGGTTTCACTCCAGGGAGTTCCACAGAAA 677  
QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87  
Db 678 TACCGATCCAGCTGATGTAGACCTCTCACCATTAATCTATCTCTGTCATCTGATGG 737  
QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107  
Db 738 GTCTCTACTGTGAATGACCAAGAGAAACAG-----GTCTCTGCGCTGAGCGC 785

QY 108 AlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121  
Db 786 ACCATTCCCATCACCCTGGAGAGAACCTTGCTGTCCAGCGAGCCCC 833  
RESULT 10  
US-09-513-999C-3892  
; Sequence 3892, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 3892  
; LENGTH: 695  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 269..592  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 44  
; OTHER INFORMATION: s=g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 45  
; OTHER INFORMATION: s=g or c  
US-09-513-999C-3892  
Alignment Scores:  
Pred. No.: 4.19e-37 Length: 695  
Score: 330.00 Matches: 62  
Percent Similarity: 77.27% Conservative: 23  
Best Local Similarity: 56.36% Mismatches: 19  
Query Match: 51.64% Indels: 6  
DB: 4 Gaps: 2  
US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-513-999C-3892 (1-695)  
QY 14 ValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGlu 33  
Db 269 ATCGGCTGGAGAGGACAGGTTCTCTGTCAACTGGATGTGAGTGCATGGAAAAATCCAGATCCAGCTGAT 328  
QY 34 AspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArg 53  
Db 329 GAACTCAAGTTAAGGTGTGGAGATGTGATGAGTGCATGGAAAAATCCAGATCCAGCTGAT 388  
QY 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgTyrArgLeuProSerAsn 73  
Db 389 CAGGATGAACATGGTTTCACTCCAGGGAGTTCCACAGGAAATACCGATCCAGCTGAT 448  
QY 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 93  
Db 449 GTAGACCTCTCACCATTAATCTATCTGTATGGGTCTCTACTGTGATGGA 508  
QY 94 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113  
Db 509 CCAAGGAAACAG-----GTCTCTGCGCTGAGCGCAGCACCATTCCCATCCCGT 556  
QY 114 GluGluLysPro-----ThrSerAlaPro 121  
Db 557 GAAGAGAACCTTGCTGTCCAGCGAGCCCCC 586



RESULT 11  
 US-09-621-976-3172  
 ; Sequence 3172, Application US/09621976  
 ; Patent No. 6639063  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Jobert, S.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET.054PR2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21  
 ; NUMBER OF SEQ ID NOS: 19335  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 3172  
 ; LENGTH: 826  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 557..736  
 US-09-621-976-3172

Alignment Scores:  
 Pred. No.: 5 23e-34 Length: 826  
 Score: 309.00 Matches: 63  
 Percent Similarity: 76.99% Conservative: 24  
 Best Local Similarity: 55.75% Mismatches: 19  
 Query Match: 48.36% Indels: 7  
 DB: 4 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-621-976-3172 (1-826)

Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27  
 Db 454 GACACTGACCTCTCAGAGATGCGCTGGAGAGACAGGTTCTCTGTCACCTGGATGG 513  
 Qy 28 LysHisPheSerProGluAspLeuThrValIleValGlnAspPhe-ValGluIleHi 47  
 Db 514 AAGCACTTCTCCCGACAGAGAACTCAAGTTAAGGTGTTGGGAGATGTCGATTGAGGTGCA 573  
 Qy 47 sGlyIleHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgAr 67  
 Db 574 TGGAAACATGAAGAGCCCGAGGATGAACATGTTTCATCTCCAGGAGTCCACAGAAA 633  
 Qy 67 gTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspG1 87  
 Db 634 ATACGGATCCAGCTGATGTAGACCTCTCACCAATTACTTCTATCCCTGTCTATCTGATGG 693  
 Qy 87 yMetLeuThrPheCysGlyProIleGlnThrGlyLeuAspAlaThr-HisAlaGluA 107  
 Db 694 GGTCTCACTGTGAATGGACCAAGGAAACAG-----GTCTCTGGGCCCTGAGC 741  
 Qy 107 rGAlaIleProValSer-ArgGluGluysPro 117  
 Db 742 GCACCATTCCTCATCACCCTCGTGAAGAGAGCCT 774

RESULT 12  
 US-09-623-655-1300  
 ; Sequence 1300, Application US/09023655  
 ; Patent No. 6607879  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cocks, Benjamin G.  
 ; APPLICANT: Susan G. Stuart  
 ; APPLICANT: Jeffrey J. Seilhamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
 ; TITLE OF INVENTION: EXPRESSION  
 ; NUMBER OF SEQUENCES: 1508  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA

COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/023,655  
 FILING DATE: HEREWITH  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0001 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1300:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 789 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: g32477  
 US-09-023-655-1300

Alignment Scores:  
 Pred. No.: 2.72e-27 Length: 789  
 Score: 261.50 Matches: 55  
 Percent Similarity: 65.18% Conservative: 18  
 Best Local Similarity: 49.11% Mismatches: 30  
 Query Match: 40.92% Indels: 10  
 DB: 4 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-023-655-1300 (1-789)

Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 293 CTCAGACGCGGGTCTCGAGATCCGCACACTGCGACCGCTGGCGGTGTCTCTGGAT 352  
 Qy 27 ValIleHisPheSerProGluAspLeuThrValIleValGlnAspAspPheValGluIle 46  
 Db 353 GTCAACCACTTTCGCCCCGCGACGCTGACGCTCAAGACCAAGGATGGCGTGTGGAGATC 412  
 Qy 47 HisGlyIleHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
 Db 413 ACCGGCAGCAGCAGGAGCGGACGACGAGATGGTACATCTCCCGGTGCTTCACGGCG 472  
 Qy 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 473 AAATACAGCTGCCCCCGGTGTGGAGCCCAAGTTTCTCTCTCTCTCTCTCTCTCTGAG 532  
 Qy 87 GlyMetLeuThrPheCysGlyProIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 533 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 564  
 Qy 107 ArgAlaIleProValSerArgGluGluysProThr 118  
 Db 565 GCTAGCAGCGAGTCCCAACGAGATCACCATCCCACT 600

RESULT 13  
 US-09-949-016-4992  
 ; Sequence 4992, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.

;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
;; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
;; FILE REFERENCE: CL001307

;; CURRENT APPLICATION NUMBER: US/09/949,016

;; CURRENT FILING DATE: 2000-04-14

;; PRIOR APPLICATION NUMBER: 60/241,755

;; PRIOR FILING DATE: 2000-10-20

;; PRIOR APPLICATION NUMBER: 60/237,768

;; PRIOR FILING DATE: 2000-10-03

;; PRIOR APPLICATION NUMBER: 60/231,498

;; PRIOR FILING DATE: 2000-09-08

;; NUMBER OF SEQ ID NOS: 207012

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 4992

;; LENGTH: 845

;; TYPE: DNA

;; ORGANISM: Human

US-09-949-016-4992

Alignment Scores:

Pred. No.: 3,02e-27 Length: 845  
Score: 261.50 Matches: 55  
Percent Similarity: 65.18% Conservative: 18  
Best Local Similarity: 49.11% Mismatches: 30  
Query Match: 40.92% Indels: 10  
DB: 4 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-949-016-4992 (1-845)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
Db 348 CTCAGCAGCGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGGTGCTCGGAT 407  
QY 27 VallysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLe 46  
Db 408 GTCAACCACTTCGCCCGGACGAGCTGACGCTCAGACCAAGATGGCGTGTGAGATC 467  
QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrlleSerArgGluPheHisArg 66  
Db 468 ACCGGCAAGCAGCAGGAGCGGACGAGCAGCATGCTACATCTCCGGTGTTCACGCGG 527  
QY 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 528 AAATACACGCTGCCCGCGTGGACCCCAAGTTTCCTCCTCCCTGCTCCCTGAG 587  
QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
Db 588 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 619  
QY 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
Db 620 GCTAGCCAGCGAGTCCACGAGATCACCATCCCACT 655

RESULT 14

US-09-513-783A-169

;; Sequence 169, Application US/09513783A

;; Patent No. 6416959

;; GENERAL INFORMATION:

;; APPLICANT: Giuliano, Kenneth A.

;; APPLICANT: Kapur, Ravi

;; TITLE OF INVENTION: A System for Cell Based Screening

;; FILE REFERENCE: 97-022-L1

;; CURRENT APPLICATION NUMBER: US/09/513,783A

;; CURRENT FILING DATE: 2000-02-25

;; NUMBER OF SEQ ID NOS: 180

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 169

;; LENGTH: 1380

;; TYPE: DNA

;; ORGANISM: Artificial Sequence

;; FEATURE:

;; OTHER INFORMATION: Description of Artificial Sequence: GFP-HSP27

;; NAME/KEY: CDS

;; LOCATION: (1)..(1380)

US-09-513-783A-169

Alignment Scores:

Pred. No.: 6,37e-27 Length: 1380  
Score: 261.50 Matches: 55  
Percent Similarity: 65.18% Conservative: 18  
Best Local Similarity: 49.11% Mismatches: 30  
Query Match: 40.92% Indels: 10  
DB: 3 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-513-783A-169 (1-1380)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
Db 1003 CTCAGCAGCGGGTCTCGGAGATCCGGCACACTGCGGACCGCTGGCGGTGCTCGGAT 1062  
QY 27 VallysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLe 46  
Db 1063 GTCAACCACTTCGCCCGGACGAGCTGACGCTCAGACCAAGGATGGCGTGTGAGATC 1122  
QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrlleSerArgGluPheHisArg 66  
Db 1123 ACCGGCAAGCAGCAGGAGCGGACGAGCATGCTACATCTCCCGGTGCTTACGCGG 1182  
QY 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 1183 AAATACACGCTGCCCGCGTGGACCCCAAGTTTCCTCCTCCCTGCTCCCTGAG 1242  
QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
Db 1243 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 1274  
QY 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
Db 1275 GCTAGCCAGCGATCCACGAGATCACCATCCCACT 1310

RESULT 15

US-09-553-498-5

;; Sequence 5, Application US/09553498

;; Patent No. 6309861

;; GENERAL INFORMATION:

;; APPLICANT: Ambrosius, Dorothee

;; APPLICANT: Rudolph, Rainer

;; APPLICANT: Schaeffner, Joerg

;; APPLICANT: Schwarz, Elisabeth

;; TITLE OF INVENTION: Process for the production of naturally folded and secreted prote

;; FILE REFERENCE: Case 20379

;; CURRENT APPLICATION NUMBER: US/09/553,498

;; PRIOR FILING DATE: 2000-04-20

;; PRIOR FILING DATE: 1999-04-26

;; NUMBER OF SEQ ID NOS: 10

;; SEQ ID NO 5

;; LENGTH: 1379

;; TYPE: DNA

;; ORGANISM: E. coli

;; FEATURE:

;; NAME/KEY: CDS

;; LOCATION: (392)...(1090)

US-09-553-498-5

Alignment Scores:

Pred. No.: 1,04e-26 Length: 1379  
Score: 260.00 Matches: 57  
Percent Similarity: 65.18% Conservative: 16  
Best Local Similarity: 50.89% Mismatches: 33  
Query Match: 40.69% Indels: 6  
DB: 3 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-553-498-5 (1-1379)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26

Db 713 CTCACGAGCGGGTCTCGGAGATCCGACAGCGGTGATCGTGGCGGTGCTCCCTGGAC 772  
Qy 27 VallysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluile 46  
Db 773 GTCAACCACTTCGCTCCGAGGAGCTCACAGTGAAGCAAGAGCGGTGGAGATC 832  
Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 66  
Db 833 ACTGGAACGACGAAGAAAGGAGGAGGAGCAACATGCTACTCTCGGTGCTTCACCCGG 892  
Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 893 AAATACACGCTCCCTCCAGGTGTGGACCCCACTAGTGTCTCTCCCTATCCCTGAG 952  
Qy 87 GlyMetLeuThrPheCysGly-----ProLysIleGlnThrGlyLeuAspAlaThrHis 104  
Db 953 GGCACACTTACCGTGGAGGCTCGGTGTCGCAAGCAGTCACG-----CAGTCA 1000  
Qy 105 AlaGluArgAlaIleProValSerArgGluGluLys 116  
Db 1001 GCGGAGATCACCATTCCGGTTACTTTTCGAGGCCCGC 1036

## RESULT 16

US-09-618-869-5  
; Sequence 5, Application US/09618869  
; Patent No. 6455279  
; GENERAL INFORMATION:  
; APPLICANT: Ambrosius, Dorthoe  
; APPLICANT: Rudolph, Rainer  
; APPLICANT: Schaeffner, Joerg  
; APPLICANT: Schwarz, Elisabeth  
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND  
; TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR  
; TITLE OF INVENTION: CHAPERONES  
; FILE REFERENCE: 20381  
; CURRENT APPLICATION NUMBER: US/09/618,869  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: EP99114811.5  
; PRIOR FILING DATE: 1999-07-29  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1379  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (392)..(1090)  
US-09-618-869-5

Alignment Scores:  
Pred. No.: 1,04e-26 Length: 1379  
Score: 260.00 Matches: 57  
Percent Similarity: 65.18% Conservatives: 16  
Best Local Similarity: 50.89% Mismatches: 33  
Query Match: 40.69% Indels: 2  
Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-618-869-5 (1-1379)

Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
Db 713 CTCACGAGCGGGTCTCGGAGATCCGACAGCGGTGATCGTGGCGGTGCTCCCTGGAC 772  
Qy 27 VallysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluile 46  
Db 773 GTCAACCACTTCGCTCCGAGGAGCTCACAGTGAAGCAAGAGCGGTGGAGATC 832  
Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 66  
Db 833 ACTGGAACGACGAAGAAAGGAGGAGGAGCAACATGCTACTCTCGGTGCTTCACCCGG 892

Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 893 AAATACACGCTCCCTCCAGGTGTGGACCCCACTAGTGTCTCTCCCTATCCCTGAG 952  
Qy 87 GlyMetLeuThrPheCysGly-----ProLysIleGlnThrGlyLeuAspAlaThrHis 104  
Db 953 GGCACACTTACCGTGGAGGCTCGGTGTCGCAAGCAGTCACG-----CAGTCA 1000  
Qy 105 AlaGluArgAlaIleProValSerArgGluGluLys 116  
Db 1001 GCGGAGATCACCATTCCGGTTACTTTTCGAGGCCCGC 1036

## RESULT 17

US-09-949-016-14287  
; Sequence 14287, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14287  
; LENGTH: 7125  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(7125)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14287

Alignment Scores:  
Pred. No.: 2.13e-15 Length: 7125  
Score: 188.00 Matches: 37  
Percent Similarity: 75.36% Conservatives: 15  
Best Local Similarity: 53.62% Mismatches: 11  
Query Match: 29.42% Indels: 6  
Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-949-016-14287 (1-7125)

Qy 55 AspAspHisGlyTyrIleSerArgGluPheHisArgTyrArgLeuProSerAsnVal 74  
Db 4783 GATGAACATGTTTCATCTCCAGGAGTTCACAGAAATACCGATCCACCTGATGTA 4842  
Qy 75 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyPro 94  
Db 4843 GACCCTCTCACCATTACTTCATCTCCCTGATGCGGTCTCTCCTGATGATGACCA 4902  
Qy 95 LysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGlu 114  
Db 4903 AGGAACAG-----GTCTCTGCGCTGAGCGGACCATTCCTCCATCCCGTGAA 4950  
Qy 115 GluLysPro-----ThrSerAlaPro 121  
Db 4951 GAGAAGCTGCTGTCCACCGACGCCCC 4977

## RESULT 18

US-09-513-999C-10503  
; Sequence 10503, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.

```
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36881
; SOFTWARE: Patent.Pm
; SEQ ID NO 10503
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 75
; OTHER INFORMATION: s=g or c
US-09-513-999C-10503

Alignment Scores:
Pred. No.: 1.57e-16 Length: 393
Score: 182.50 Matches: 42
Percent Similarity: 60.23% Conservative: 11
Best Local Similarity: 47.73% Mismatches: 26
Query Match: 28.56% Indels: 10
DB: 4 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-09-513-999C-10503 (1-393)
QY 31 SerProGluAspLeuThrValLysValGlnAspPheValGluIleHisGlyLysHis 50
Db 36 AGCATGAACGAGTGTGAGGTCAAGACCAAGAGTGGTGTGGAGATCACCGGCAAGCAC 95
QY 51 AsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArgTyrArgLeu 70
Db 96 GAGGAGCGGCGAGACGAGTGTGATCTCCCGTGTCTTCCGCGGAATACAGCGTGT 155
QY 71 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 90
Db 156 CCCCCCGGTGTGACCCGCCACCCAGTTTCTCTCCCTGTCCCTGAGGGCACACTGACC 215
QY 91 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 110
Db 216 GTGGA-GGCCCC-----CATGCCAAGCTAGCCACGCA 247
QY 111 ValSerArgGluGluLysProThr 118
Db 248 GTCCACGAGATCACCATCCCACT 271

RESULT 19
US-08-900-407-2
; Sequence 2, Application US/08900407
; Patent No. 5962262
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: NEW HUMAN HEAT SHOCK 27 LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,407
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0351 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1627 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE: 1
; INDIVIDUAL ISOLATE:
; IMMEDIATE SOURCE:
; LIBRARY: 22??
; CLONE: 1362715
US-08-900-407-2

Alignment Scores:
Pred. No.: 2.96e-12 Length: 1627
Score: 159.00 Matches: 28
Percent Similarity: 64.86% Conservative: 20
Best Local Similarity: 37.84% Mismatches: 26
Query Match: 24.88% Indels: 0
DB: 2 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-08-900-407-2 (1-1627)
QY 23 IlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGlnAspAsp 42
Db 808 GTGTGTGTGAATGTGCACAGCTTCAAGCCAGAGAGTGTGATGGTGAAGACCAAGATGGA 867
QY 43 PheValGluIleHisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArg 62
Db 868 TACGTGAGGTGTCTGCAACATGAAGAAACAGCAAGAGGTGGCATTTGTTTCTAAG 927
QY 63 GluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSer 82
Db 928 AACTTCACAAAGAAATCCAGCTTCTCTGCAGAGTGGATCTCTGTGACAGTATTTGCTCA 987
QY 83 LeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIle 96
Db 988 CTTTCCCAGAGGTGTGCTGATCATCGAAGCTCCCGAGTGC 1029

RESULT 20
US-09-949-016-91317
; Sequence 91317, Application US/09949016
; Patent No. 8812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```





```
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15845

Alignment Scores:
Pred. No.: 7.14e-08 Length: 599
Score: 123.50 Matches: 28
Percent Similarity: 52.94% Conservative: 17
Best Local Similarity: 32.94% Mismatches: 39
Query Match: 19.33% Indels: 1
DB: 4 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-09-270-767-15845 (1-599)

QY 11 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 30
Db 260 CTAAGTCGGGTAAATTCGGAAGGATGTTTCGAGGCAATGTGATGTCACCTGTTTC 319
QY 31 SerProGluAspLeuThrValLysValGlnAspPheValGluLeuHisGlyLysHis 50
Db 320 AAGCCCTATGAGATTAGCTGGAAGACCTCAGCGCACACTGTGCTGCGAGGCCAAGCAC 379
QY 51 AsnGluArgGlnAspAspHisGlyTyrlleSerArgGluPheHisArgArgTyrArgLeu 70
Db 380 GAGAACGACGATGATGTCACACCTTCGTGGTGGCCACATCGTCAAGCGGTTGCTCTG 439
QY 71 ProSerAsnValAspGlnSerAlaLeuSerCysSerSerAlaAspGlyMetLeuThr 90
Db 440 CCCCAGGAGTACTATCCCAACGATGTCGATCGGAACCTGCTGCGCATGCTACTTACC 499
QY 91 Phe---CysGlyPro 94
Db 500 GTCAAGTGTCCGCCG 514

RESULT 27
US-09-949-016-16734
; Sequence 16734, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16734
; LENGTH: 5688
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16734

Alignment Scores:
Pred. No.: 9.56e-06 Length: 5688
Score: 119.00 Matches: 22
Percent Similarity: 74.42% Conservative: 10
Best Local Similarity: 51.16% Mismatches: 11
Query Match: 18.62% Indels: 0
DB: 4 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-09-949-016-16734 (1-5688)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26
Db 2348 CTCAGCAGCGGGGTCTCGAGATCCGGCACACTGCGGACCGCTGCGCGTCCCTGGAT 2407

QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluLeu 46
Db 2408 GTCAACCACTTCGCCCCGCGAGCTGACGGTCAGAGCAAGCATGCGGTGGTGAGATC 2467
QY 47 HisGlyLys 49
Db 2468 ACCGGTGAG 2476

RESULT 28
US-09-669-751-69
; Sequence 69, Application US/09669751
; Patent No. 6551575
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; Motion Sickness, Vertigo and Other Disorders Related to
; Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/09/669,751
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Drosophila
US-09-669-751-69

Alignment Scores:
Pred. No.: 3.23e-07 Length: 495
Score: 118.00 Matches: 21
Percent Similarity: 68.29% Conservative: 7
Best Local Similarity: 51.22% Mismatches: 13
Query Match: 18.47% Indels: 0
DB: 4 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-09-669-751-69 (1-495)

QY 18 ArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrVal 37
Db 368 AAAGATGGCTTCACGTGGCATGGATGTGCGAGTTCAGCCCAACGAGCTGACCGTC 427
QY 38 LysValGlnAspAspPheValGluLeuHisGlyLysHisAsnGluArgGlnAspAspHis 57
Db 428 AAGGTGGTGACACACACCTGCTGTAGAGGGCCAGCACGAGGAGCGGAGCGCCAT 487
QY 58 Gly 58
Db 488 GGA 490

RESULT 29
US-09-949-016-89648/c
; Sequence 89648, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89648
; LENGTH: 601
```



```

; TYPE: DNA
; ORGANISM: Human
US-09-949-016-89648

Alignment Scores:
Pred. No.: 4.33e-07 Length: 601
Score: 118.00 Matches: 21
Percent Similarity: 82.76% Conservative: 3
Best Local Similarity: 72.41% Mismatches: 5
Query Match: 18.47% Indels: 0
DB: 4 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-09-949-016-89648 (1-601)

Qy 26 AspVallyshisPheSerProGluAaspLeuThrVallyValGlnAaspPheValGlu 45
Db 601 GATGTGAAGCATTCTCCCGAGGAACTCAAGTTAAGGTGTGGAGATGTGATTGAG 542
Qy 46 IleHisGlylyshisAsnGluArgGln 54
Db 541 GTGCATGGAACATGAGAGGCCAG 515

RESULT 30
US-09-949-016-17284/c
; Sequence 17284, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17284
; LENGTH: 12214
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17284

Alignment Scores:
Pred. No.: 0.000113 Length: 12214
Score: 115.00 Matches: 35
Percent Similarity: 51.35% Conservative: 22
Best Local Similarity: 31.53% Mismatches: 36
Query Match: 18.00% Indels: 18
DB: 4 Gaps: 5

US-10-657-740-1_COPY_51_173 (1-123) x US-09-949-016-17284 (1-12214)

Qy 4 ArgThrValLeuAaspSerGlyIleSerGluValArgSerAasp-----ArgAaspysphe 21
Db 376 CGGCTGCTCAGGACAGTCCAGCGCTCAGGAGGACATGACATGCCAGAGCGGTTTC 317
Qy 22 ValIlePheLeuAaspVallyshisPheSerProGluAaspLeuThrVallyValGlnAasp 41
Db 316 CAATGAGCTGGATGCCCGACCGCTTCGCCCGGAGAACTGGTGGTGCAGTGCATGGC 257
Qy 42 AspPheValGluIleHisGlylyshisAsn-----GluArgGlnAasp 55
Db 256 CAATGGCTGATGTCACCGGACAGCAGCAACTGGACGTGAGGAGCCCGAAAGGGTCAGT 197
Qy 56 AspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAasp 75
Db 196 TACCCG-----ATGTACAGAAAGGTGACCGGAAA---ATGCTACCGTCCAACTGAGT 146

```

```

Qy 76 GlnSerAlaLeuSerCysSerLeuSerAlaAaspGlyMetLeuThrPheCysGly----- 93
Db 145 CCTACCGCCATGACCTGCTGCTGACCCCTCGGCGCAGCTGTGGGTGAGGCCAGTGT 86
Qy 94 -----ProlyslleGlnThrGly 99
Db 85 GTGGCGTGGCCCTCCCTGAAGCCCAACAGGA 53

RESULT 31
US-09-620-312D-669
; Sequence 669, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Aseundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-bong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, fonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radjoe T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 669
; LENGTH: 719
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(578)
US-09-620-312D-669

Alignment Scores:
Pred. No.: 6.62e-06 Length: 719
Score: 110.50 Matches: 23
Percent Similarity: 54.32% Conservative: 21
Best Local Similarity: 28.40% Mismatches: 34
Query Match: 17.29% Indels: 3
DB: 4 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-09-620-312D-669 (1-719)

Qy 10 GlyIleSerGluValArgSerAaspysPheValIlePheLeuAaspVallyshis 29
Db 276 GGGGCGAGCAACATCAAGACCTAGGACGCCCTATGAGTTTGGGTGGACGTGAGAGAC 335
Qy 30 PheSerProGluAaspLeuThrVallyValGlnAaspPheValGluIleHisGlylys 49
Db 336 TTCTCCTCTGAGACATCATTTGTCCACCCTCCAAACACCATCGAGTGGCGGCT--- 392
Qy 50 HisAsnGluArgGlnAaspPheHisGlyTyrIleSerArgGluPheHisArgArgTyrArg 69
Db 393 -----GAGAAGCTGGCGGCTGACGGCACTGTCTATGAACACCTTCGCTCACAAGTCCAG 446
Qy 70 LeuProSerAsnValAaspGlnSerAlaLeuSerCysSerLeuSerAlaAaspGlyMetLeu 89

```

```
Db 447 CTGCGGAGGAGTGGACCCGACGTCGGTCTGCGGGAGGACGGACGCTC 506
Qy 90 Thr 90
Db 507 ACT 509

RESULT 32
US-09-949-016-17282/c
; Sequence 17282, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17282
; LENGTH: 9615
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17282

Alignment Scores:
Pred. No.: 0.000403 Length: 9615
Score: 110.00 Matches: 31
Percent Similarity: 51.58% Conservative: 18
Best Local Similarity: 32.63% Mismatches: 30
Query Match: 17.21% Indels: 16
DB: 4 Gaps: 4

US-10-657-740-1_COPY_51_173 (1-123) x US-09-949-016-17282 (1-9615)
Qy 18 ArgAspLysPheValIlePheLeuAspValLysPheSerProGluAspLeuThrVal 37
Db 9588 AGAGACGGTTTCCAAATGAAGCTGGATGCCACCGCTTCCGCCCGGAGGAACTGGTG 9529
Qy 38 LysValGlnAspAspPheValGluLeuHisGlyLysHisAsn----- 51
Db 9528 CAGGTGGATGCCAATGGCTGATGGTACCGCAGCAGCACTGGAGCTCAGGACCCG 9469
Qy 52 GluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 71
Db 9468 GAAAGGGTCAGTACCAGC-----ATGTCACAGAGGTGCACCGGAAA--ATGCTACCG 9418
Qy 72 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 91
Db 9417 TCCAACTGAGTCTCTACCGCATGACCTGCTGCTGACCCCTCCGGCGCAGCTGGGTC 9358
Qy 92 CysGly-----ProLysIleGlnThrGly 99
Db 9357 AGAGGCCAGTGTGTGGCGCTGGCCCTCTCCCTGAAGCCCAACAGGA 9313

RESULT 33
US-09-949-016-14287/c
; Sequence 14287, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14287
; LENGTH: 7125
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(7125)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14287

Alignment Scores:
Pred. No.: 0.000682 Length: 7125
Score: 107.00 Matches: 18
Percent Similarity: 69.05% Conservative: 11
Best Local Similarity: 42.86% Mismatches: 13
Query Match: 16.74% Indels: 0
DB: 4 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-09-949-016-14287 (1-7125)
Qy 9 SerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLys 28
Db 130 GCAGGGCCCTCCGAGCTTAGCTCAGTCAGGAGGCGAGTTCCAGGCATTTCTGGATGTGAGC 71
Qy 29 HisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLeuHisGly 48
Db 70 CACTTTACCCACAGCAGGAGTGACTGTGAGGACTGTGGATAACCTGCTGGAGGTGTCTGCC 11
Qy 49 LysHis 50
Db 10 CGGCAC 5

RESULT 34
US-09-949-016-175645
; Sequence 175645, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 175645
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-175645

Alignment Scores:
Pred. No.: 0.000417 Length: 601
Score: 97.00 Matches: 29
Percent Similarity: 39.33% Conservative: 6
Best Local Similarity: 32.58% Mismatches: 13
Query Match: 15.18% Indels: 41
DB: 4 Gaps: 2
```

RESULT 35  
 US-09-107-532A-670  
 ; Sequence 670, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD-ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107,532A  
 ; FILING DATE: 30-Jun-1998  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085,598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ariniello, Pamela Deneke  
 ; REGISTRATION NUMBER: 40,489  
 ; REFERENCE/DOCKET NUMBER: GTC-012  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (781)893-5007  
 ; TELEFAX: (781)893-8277  
 ; INFORMATION FOR SEQ ID NO: 670:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 420 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: circular  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Enterococcus faecium  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (B) LOCATION 1..420  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 670:

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 861-3000  
 TELEFAX: (202) 822-0944  
 TELEX: 6714627 CUSH  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 652 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 US-07-749-761-2

Alignment Scores:		
Pred. No.:	0.0076	Length:
Score:	88.50	Matches:
Percent Similarity:	51.22%	Conservative:
Best Local Similarity:	30.49%	Mismatches:
Query Match:	13.85%	Indels:
DB:	1	Gaps:

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-07-748-761-2 (1-652)

Qy	22	ValliePheLeuAspValIlyHisPheSerProGluAspLeuThrVallysValGlnAsp	41
Dd	219	GTTGTTCAAAGTTCATCTCCACGGCTTAAGAAGCAGGAAGTCAAAGTGCAGATCGAGGAG	278
Qy	42	AspPheVal--GluileHiscglyLylYHisEandGluArgGlnAspAspHisGly----	58
Dd	279	GATAGGTTCTTCAGTTCAGCGAAGGAGGAACTGGAGGAGGAATAAGAATGATAG	338
Qy	59	-----TyrIlSerArgGluPheHisArgArgTyArgLeuProSerAsn	73
Dd	339	TGGCATCGCATGGACGGAAGCAGCGGGAATTCA TGAGGAGATTTAGACTTCGCGGAAT	398
Qy	74	ValAspGlnSerAlaLeuSerCyseSerleuSerAlaaspGlyMetLeuThrPheCySGly	93
Dd	399	GC AAAGATGGATCAAGTTAAGCGGTCTATGGAG---AATGGAGTGCTTACTGTACTGTT	455
Qy	94	ProLys	95
Dd	456	CCAAAG	461

```

RESULT 37
US-09-949-016-13807/c
; Sequence 13807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13807
; LENGTH: 168971
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(168971)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13807

```

**Alignment Scores:**

[illegible]

**Alignment Scores:**

Alignment Scores:					
Pred. No.:	0.0111	Length:	489		
Score:	86.00	Matches:	33		
Percent Similarity:	43.3%	Conservative:	16		
Best Local Similarity:	29.2%	Mismatches:	39		
Query Match:	13.4%	Indels:	26		
DB:	4	Gaps:	4		

US-10-657-740-1 COPY 51 173 (1-123) x US-09-902-540-6436 (1-489)

Qy.		' 13	GluValArgSerAspArgAspLysPheValIlePheLeuAspVallylHisPheSerPro	32
			:::::	
			:::::	
			:::::	
			:::::	
Db		154	GAGGTGAAGAGACCCAGGATGCCTTCATCTTCAAGCGCGATGTCGCCGCGTGAGGAG	213
			:::::	
			:::::	
			:::::	
			:::::	
Qy		33	GluAspLeuThrVallylsvalGlnAspAspPheValGluIleHisGlyLysHisAsnGlu	52
			:::	
			:::	
			:::	
			:::	
Db		214	AAGGACCTGGAAATCACCTCGCGAGAACCGGCTGCACCATCAGCGCAAGCGGGAAGAA	273
			:::::	
			:::::	
			:::::	
			:::::	
			:::::	
Qy		53	ArgGlnAspAspHisGly-----TyrIleSerHrgLu-----PheHis	65
			:::	
			:::	
			:::	
			:::	
			:::	
Db		274	GAGCGCCGCGAAGGCGCGCGGTACTACGCCCTACGAGCGCAACTATGCTGCTGTCAGC	333
			:::	
			:::	
			:::	
			:::	
			:::	

**Alignment Scores:**

```

Qy 66 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 85
Db 334 CGACCTTCAGCTCCCGGGGTGTG-----AATGCC 366

Qy 86 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 105
Db 367 GAC-----AAGCTCCAGGC-GGACTTCAAGAGTGGCGTCCT 401

Qy 106 GluArgAla-IleProValSerArgGluGluLysPro 117
Db 402 CAACGTGGCATCCCCAAGAAGTCCGAGGAGCAGCCC 438

RESULT 39
US-09-902-540-444
; Sequence 444, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 444
; LENGTH: 3626
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(3626)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-444

Alignment Scores:
Pred. No.: 0.234 Length: 3626
Score: 86.00 Matches: 33
Percent Similarity: 43.36% Conservative: 16
Best Local Similarity: 29.20% Mismatches: 39
Query Match: 13.46% Indels: 26
DB: 4 Gaps: 4

US-10-657-740-1_COPY_51_173 (1-123) x US-09-902-540-444 (1-3626)
Qy 13 GluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSerPro 32
Db 1509 GAGGTGAGGAGACCCAGGATCCCTTCATCTTCAAGCGCGGATGTCCTCCCGGGTGGAGGAG 1568

Qy 33 GluAspLeuThrValLysValGlnAspPheValGluIleHisGlyLysHisAsnGlu 52
Db 1569 AAGGACCTTGAATCACCTCGCGGAAACCGGCTGACCATCAGCGGCAACGGGAAGAA 1628

Qy 53 ArgGlnAspAspHisGly-----TyrIleSerArgGlu-----PheHis 65
Db 1629 GAGCGCCGCGAGCAAGAGGACCGCTACTACGCTACGAGCGCAACTATGGCTGCTTCAGC 1688

Qy 66 ArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAla 85
Db 1689 CGACCTTCAGCTCCCGGGGTGTG-----AATGCC 1721

Qy 86 AspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAla 105
Db 1722 GAC-----AAGCTCCAGGC-GGACTTCAAGAGTGGCGTCCT 1756

Qy 106 GluArgAla-IleProValSerArgGluGluLysPro 117
Db 1757 CAACGTGGCATCCCCAAGAAGTCCGAGGAGCAGCCC 1793

RESULT 40

```

```

US-09-489-039A-3463
; Sequence 3463, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3463
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3463

Alignment Scores:
Pred. No.: 0.0198 Length: 576
Score: 85.00 Matches: 23
Percent Similarity: 40.00% Conservative: 19
Best Local Similarity: 21.90% Mismatches: 43
Query Match: 13.30% Indels: 20
DB: 4 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x US-09-489-039A-3463 (1-576)
Qy 17 AspArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThr 36
Db 289 GACGAAACCACTATCGCATCGCTATCGCGTGGCTTGTGTAACGCGAGCTGGAG 348

Qy 37 ValLysValGlnAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAsp 56
Db 349 ATCACTGCCAGGACCAATCTGCTGATCGTCAAAGGCGCCACCGCCGAGCAGAGAG 408

Qy 57 HisGlyTyrIle-----SerArgGluPheHisArgTyrArgLeuPro 71
Db 409 CGGACCTTACCTGTATCAGGGGATCGCGGAGCGCAACTTCGAGCGCAATTCAGCTGGCG 468

Qy 72 SerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 91
Db 469 GAAACACATTCACCTCCGC-----486

Qy 92 CysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProVal 111
Db 487 ---GGCGCCAATCTGTTGAACGCGCTGTATATCGATCTGAAACGGGTGATCCCGGAA 543

Qy 112 SerArgGluGluLys 116
Db 544 GCGAACACGCGCGC 558

RESULT 41
US-09-583-110-2262
; Sequence 2262, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2262
; LENGTH: 1350

```

```
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2262

Alignment Scores:
Pred. No.: 0.437 Length: 1350
Score: 79.50 Matches: 25
Percent Similarity: 46.99% Conservative: 14
Best Local Similarity: 30.12% Mismatches: 27
Query Match: 12.44% Indels: 17
DB: 4 Gaps: 4

US-10-657-740-1_COPY_51_173 (1-123) x US-09-583-110-2262 (1-1350)
QY 26 AspVallyHisPheSerProGluAspLeuThrValGln---AspAspPheVal 44
DB 67 GACATCAAGGCCAACAGTCAAGTAAATGTTGATGCTTTTCAAGAGCTGATCAGTTG 126
QY 45 GluIleHisGlyLyHisAsnGluArgGlnAspAspHisGlyTyr----- 59
DB 127 AGTATTACCTTAAGCAACAGCATTTTCCCAAGAGCATGAACCTCTTAAAGATATCAAG 186
QY 60 IleSerArgGluPheHisArgGlyTyrArgLeuProSerAsnValAsp----- 75
DB 187 ATCGGCAACACTCTTTTAAAGAGGTGAGCTCCCTTCTAAGTTGATTCAGTCGTAAAA 246
QY 76 -----GlnSerAlaLeuSerCysSerLeuSerAla-----AspGlyMet 88
DB 247 GTTACTTTGAAGAGTGTAGTGTCTTCTAGCCAGGCGATGCTGCGCGCATG 306
QY 89 LeuThrPhe 91
DB 307 GAAACTTTT 315

RESULT 42
US-09-107-433-2322
; Sequence 2322, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2322:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1353 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1353
; SEQUENCE DESCRIPTION: SEQ ID NO: 2322:
US-09-107-433-2322

Alignment Scores:
Pred. No.: 0.438 Length: 1353
Score: 79.50 Matches: 25
Percent Similarity: 46.99% Conservative: 14
Best Local Similarity: 30.12% Mismatches: 27
Query Match: 12.44% Indels: 17
DB: 4 Gaps: 4

US-10-657-740-1_COPY_51_173 (1-123) x US-09-107-433-2322 (1-1353)
QY 26 AspVallyHisPheSerProGluAspLeuThrValGln---AspAspPheVal 44
DB 70 GACATCAAGGCCAACAGTCAAGTAAATGTTGATGCTTTTCAAGAGCTGATCAGTTG 129
QY 45 GluIleHisGlyLyHisAsnGluArgGlnAspAspHisGlyTyr----- 59
DB 130 AGTATTACCTTAAGCAACAGCATTTTCCCAAGAGCATGAACCTCTTAAAGATATCAAG 189
QY 60 IleSerArgGluPheHisArgGlyTyrArgLeuProSerAsnValAsp----- 75
DB 190 ATCGGCAACACTCTTTTAAAGAGGTGAGCTCCCTTCTAAGTTGATTCAGTCGTAAAA 249
QY 76 -----GlnSerAlaLeuSerCysSerLeuSerAla-----AspGlyMet 88
DB 250 GTTACTTTGAAGAGTGTAGTGTCTTCTAGCCAGGCGATGCTGCGCGCATG 309
QY 89 LeuThrPhe 91
DB 310 GAAACTTTT 318

RESULT 43
US-09-949-016-11770/C
; Sequence 11770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11770
; LENGTH: 49487
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(49487)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11770
```

```
Alignment Scores:
Pred. No.: 202 Length: 49487
Score: 77.50 Matches: 41
Percent Similarity: 39.53% Conservative: 27
Best Local Similarity: 23.84% Mismatches: 53
Query Match: 12.13% Indels: 51
DB: 4 Gaps: 10

US-10-657-740-1_COPY_51_173 (1-123) x US-09-949-016-11770 (1-49487)
QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAsp----- 17
Db 2663 TCTCTGATTGCTAGGCGATTCTTCACCGCTTAAGTGAGCTGAGGTGTAACGACCATGAT 2604
QY 18 -----ArgAspLysPheValIlePheLeuAsp----- 26
Db 2603 AGATAATCATCCAGATCGGAAACCCCGAGAGCGCAGGATTATCTGGATAACGGAGCG 2544
QY 27 -----ValLysHisPheSerProGlu-----AspLeuThr 36
Db 2543 GATGGAGGATTGCTCGGTTCTCCACACTCAAGCCCGGAAAGGTCACCTGACCTAGTC 2484
QY 37 ValLysValGlnAspPheValGluIleHisGlyLys---HisAsn----- 51
Db 2483 TTATTCTTTACAGCAGAAATCGGGCAAAACCCATGGCGCGTCCACAACTCTATAGCCGC 2424
QY 52 -----GluArgGlnAsp-----AspHisGlyTyr 59
Db 2423 CGACCTGGGTACAGGGGAAGCGGGAACCAACACACACTTCTACGACAGCGGTTAC 2364
QY 60 IleSerArgGluPheHisArgTyrArgLeuProSerAsnVal----- 74
Db 2363 ATTAAGCGCGACGCCACACACACGCGCCACACACACACACACACACACACACACAC 2304
QY 75 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyPro 94
Db 2303 CCCACAGCTTCGTTAACTTGTGATGATAAATTAGGCGCGAGGTGACACTACTGT---CCC 2247
QY 95 LysIleGlnThr-----GlyLeuAspAlaThrHisAlaGluArgAlaIleProVal 111
Db 2246 GAACCTCAGCTGGAACGCTAGTCCAGTCACTGTCTTAAGTGCAGCTAGCAGTGCACG 2187
QY 112 SerArgGluGluLysProThrSerAlaProSerSer 123
Db 2186 TTGTCGAGGAGTTTCT---TCCGCTCCCTCACT 2154

RESULT 44
US-09-949-016-13631
; Sequence 13631, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13631
; LENGTH: 57751
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13631

Alignment Scores:
Pred. No.: 202 Length: 57751
Score: 77.50 Matches: 28
Percent Similarity: 41.84% Conservative: 13
Best Local Similarity: 28.57% Mismatches: 40
Query Match: 11.82% Indels: 18
DB: 4 Gaps: 5

US-10-657-740-1_COPY_51_173 (1-123) x US-09-949-016-13631 (1-57751)
QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAsp----- 17
Db 1009 TCTCTGATTGCTAGGCGATTCTTCACCGCTTAAGTGAGCTGAGGTGTAACGACCATG 1068
QY 18 -----ArgAspLysPheValIlePheLeuAsp----- 26
Db 1069 AGATAATCATCCAGATCGGAAACCCCGAGAGCGCAGGATTATCTGGATAACGGAGCG 1128
QY 27 -----ValLysHisPheSerProGlu-----AspLeuThr 36
Db 1129 GATGGAGGATTGCTCGGTTCTCCACACTCAAGCCCGGAAAGGTCACCTGACCTAGTC 1188
QY 37 ValLysValGlnAspPheValGluIleHisGlyLys---HisAsn----- 51
Db 1189 TTATTCTTTACAGCAGAAATCGGGCAAAACCCATGGCGCGTCCACAACTCTATAGCCGC 1248
QY 52 -----GluArgGlnAsp-----AspHisGlyTyr 59
Db 1249 CGACCTGGGTACAGGGGAAGCGGGAACCAACACACACTTCTACGACAGCGGTTAC 1308
QY 60 IleSerArgGluPheHisArgTyrArgLeuProSerAsnVal----- 74
Db 1309 ATTAAGCGCGACGCCACACACGCGCCACACACACACACACACACACACACACAC 1368
QY 75 AspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyPro 94
Db 1369 CCCACAGCTTCGTTAACTTGTGATGATAAATTAGGCGCGAGGTGACACTACTGT---CCC 1425
QY 95 LysIleGlnThr-----GlyLeuAspAlaThrHisAlaGluArgAlaIleProVal 111
Db 1426 GAACCTCAGCTGGAACGCTAGTCCAACTCACTGTCTTAAGTGCAGCTAGCAGTGCACG 1485
QY 112 SerArgGluGluLysProThrSerAlaProSerSer 123
Db 1486 TTGTCGAGGAGTTTCT---TCCGCTCCCTCACT 1518

RESULT 45
US-09-969-532-23
; Sequence 23, Application US/09969532
; Patent No. 6777232
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides Encodin
; FILE REFERENCE: LEX-0244-USA
; CURRENT APPLICATION NUMBER: US/09/969,532
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 60/237,280
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 966
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-969-532-23

Alignment Scores:
Pred. No.: 0.97 Length: 966
Score: 75.50 Matches: 28
Percent Similarity: 41.84% Conservative: 13
Best Local Similarity: 28.57% Mismatches: 40
Query Match: 11.82% Indels: 18
DB: 4 Gaps: 5
```



US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-969-532-23 (1-966)

```
QY 30 PheSerPro---GluAspLeuThrVallysValGlnAspPheVal----- 44
   |||::|||
Db 529 TTTAAACCCCTTTCGACATCAAGTGAAGTCCAGAGCTCGTTCATGTTCCCTGGGA 588
   |||::|||
QY 45 -----GluIleHisGlyLysHisAsnGluArgGlnAspHisGly--- 58
   |||::|||
Db 589 GTGCTGAGAGAGCTGAGTACCACGCAAGAAATCATTCAGGACTTTTCCCATGGAAC 648
   |||::|||
QY 59 ---TyrIleSerArgGluPheHisArgTyrArgLeuPro-----SerAsnValasp 75
   |||::|||
Db 649 ACCACAGCTTTAGTACATGATCCAGAAATAAATGCCCTACATCCAAATCTGTCA 708
   |||::|||
QY 76 GlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLys 95
   |||::|||
Db 709 TCACCTCCCAAGGACAGCACTGAGGACAACCTGGTGTCTTTGGCCATTTAGGGGGCGC 768
   |||::|||
QY 96 Ile-----GlnThrGlyLeuaspAlaThrHisAlaGluArgAlaIlePro 110
   |||::|||
Db 769 TTAGTAATGCCAAATACAGGGGT-GAGCTTACTCATACACACACGGTGCCATCCC 821
   |||::|||
```

## RESULT 46

US-09-969-532-21  
; Sequence 21, Application US/09969532  
; Patent No. 6777232  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides Encodir  
; FILE REFERENCE: LEX-0244-USA  
; CURRENT APPLICATION NUMBER: US/09/969,532  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US 60/237,280  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq For Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-969-532-21

## Alignment Scores:

Pred. No.:	1.02	Length:	999
Score:	75.50	Matches:	28
Percent Similarity:	41.84%	Conservative:	13
Best Local Similarity:	28.57%	Mismatches:	40
Query Match:	11.82%	Indels:	18
DB:	4	Gaps:	5

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-969-532-21 (1-999)

```
QY 30 PheSerPro---GluAspLeuThrVallysValGlnAspPheVal----- 44
   |||::|||
Db 562 TTTAAACCCCTTTCGACATCAAGTGAAGTCCAGAGCTCGTTCATGTTCCCTGGGA 621
   |||::|||
QY 45 -----GluIleHisGlyLysHisAsnGluArgGlnAspHisGly--- 58
   |||::|||
Db 622 GTGCTGAGAGAGCTGAGTACCACGCAAGAAATCATTCAGGACTTTTCCCATGGAAC 681
   |||::|||
QY 59 ---TyrIleSerArgGluPheHisArgTyrArgLeuPro-----SerAsnValasp 75
   |||::|||
Db 682 ACCACAGCTTTAGTACATGATCCAGAAATAAATGCCCTACATCCAAATCTGTCA 741
   |||::|||
QY 76 GlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLys 95
   |||::|||
Db 742 TCACCTCCCAAGGACAGCACTGAGGACAACCTGGTGTCTTTGGCCATTTAGGGGGCGC 801
   |||::|||
QY 96 Ile-----GlnThrGlyLeuaspAlaThrHisAlaGluArgAlaIlePro 110
   |||::|||
Db 802 TTAGTAATGCCAAATACAGGGGT-GAGCTTACTCATACACACACGGTGCCATCCC 854
   |||::|||
```

## RESULT 47

US-09-969-532-19  
; Sequence 19, Application US/09969532  
; Patent No. 6777232  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides Encodir  
; FILE REFERENCE: LEX-0244-USA  
; CURRENT APPLICATION NUMBER: US/09/969,532  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US 60/237,280  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq For Windows Version 4.0  
; SEQ ID NO 19  
; LENGTH: 1008  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-969-532-19

Alignment Scores:	1.04	Length:	1008
Pred. No.:	75.50	Matches:	28
Score:	41.84%	Conservative:	13
Percent Similarity:	28.57%	Mismatches:	40
Best Local Similarity:	11.82%	Indels:	18
Query Match:	4	Gaps:	5
DB:			

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-969-532-19 (1-1008)

```
QY 30 PheSerPro---GluAspLeuThrVallysValGlnAspPheVal----- 44
   |||::|||
Db 571 TTTAAACCCCTTTCGACATCAAGTGAAGTCCAGAGCTCGTTCATGTTCCCTGGGA 630
   |||::|||
QY 45 -----GluIleHisGlyLysHisAsnGluArgGlnAspHisGly--- 58
   |||::|||
Db 631 GTGCTGAGAGAGCTGAGTACCACGCAAGAAATCATTCAGGACTTTTCCCATGGAAC 690
   |||::|||
QY 59 ---TyrIleSerArgGluPheHisArgTyrArgLeuPro-----SerAsnValasp 75
   |||::|||
Db 691 AACACAGCTTTAGTACATGATCCAGAAATAAATGCCCTACATCCAAATCTGTCA 750
   |||::|||
QY 76 GlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLys 95
   |||::|||
Db 751 TCACCTCCCAAGGACAGCACTGAGGACAACCTGGTGTCTTTGGCCATTTAGGGGGCGC 810
   |||::|||
QY 96 Ile-----GlnThrGlyLeuaspAlaThrHisAlaGluArgAlaIlePro 110
   |||::|||
Db 811 TTAGTAATGCCAAATACAGGGGT-GAGCTTACTCATACACACACGGTGCCATCCC 863
   |||::|||
```

## RESULT 48

US-09-969-532-17  
; Sequence 17, Application US/09969532  
; Patent No. 6777232  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides Encodir  
; FILE REFERENCE: LEX-0244-USA  
; CURRENT APPLICATION NUMBER: US/09/969,532  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US 60/237,280  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq For Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-969-532-17

## Alignment Scores:

Pred. No.: 1.09 Length: 1041  
Score: 75.50 Matches: 28  
Percent Similarity: 41.84% Conservative: 13  
Best Local Similarity: 28.57% Mismatches: 40  
Query Match: 11.82% Indels: 18  
DB: 4 Gaps: 5

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-969-532-17 (1-1041)

QY 30 PheSerPro---GluAspLeuThrVallysValGlnAspPheVal----- 44  
Db 604 TTTAACCCCTTTGTCGGACATCAAGTGAAGTCCAGAGCTCGTTTCATGGTTTCCCTGGGA 663  
QY 45 -----GluIleHisGlyLysHisAsnGluArgGlnAspPheHisGly--- 58  
Db 664 GTGCTGAGAGAGCTGAGTACACGGCAGAGAAATCATTCAGGACTTTTCCCATGGAAAC 723  
QY 59 ---TyrIleSerArgGluPheHisArgTyrArgLeuPro-----SerAsnValAsp 75  
Db 724 AACACAGCTTTAGTACAAATGCATCCAGAAATAAATGCCTTACATCCAAATCTGTCA 783  
QY 76 GlnSerAlaLeuSerCysSerLeuSerAlaAspClyMetLeuThrPheCysGlyProLys 95  
Db 784 TCACCTCCCAACGACAGACTGAGGACAACTGGTGTCTTTGGCCATTTAGGGGGCGC 843  
QY 96 Ile-----GlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 110  
Db 844 TTAGTAATGCCAAATACAGGGGT-GAGCTTACTATACACACGGTGCCATCCC 896

## RESULT 49

US-09-969-532-7  
; Sequence 7, Application US/09969532  
; Patent No. 6777232  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides Encodin  
; FILE REFERENCE: LEX-0244-USA  
; CURRENT APPLICATION NUMBER: US/09/969,532  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US 60/237,280  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-969-532-7

## Alignment Scores:

Pred. No.: 2.21 Length: 1659  
Score: 75.50 Matches: 28  
Percent Similarity: 41.84% Conservative: 13  
Best Local Similarity: 28.57% Mismatches: 40  
Query Match: 11.82% Indels: 18  
DB: 4 Gaps: 5

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-969-532-7 (1-1659)

QY 30 PheSerPro---GluAspLeuThrVallysValGlnAspPheVal----- 44  
Db 1222 TTTAACCCCTTTGTCGGACATCAAGTGAAGTCCAGAGCTCGTTTCATGGTTTCCCTGGGA 1281  
QY 45 -----GluIleHisGlyLysHisAsnGluArgGlnAspPheHisGly--- 58  
Db 1282 GTGCTGAGAGAGCTGAGTACACGGCAGAGAAATCATTCAGGACTTTTCCCATGGAAAC 1341  
QY 59 ---TyrIleSerArgGluPheHisArgTyrArgLeuPro-----SerAsnValAsp 75  
Db 1342 AACACAGCTTTAGTACAAATGCATCCAGAAATAAATGCCTTACATCCAAATCTGTCA 1401

QY 76 GlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLys 95  
Db 1402 TCACCTCCCAACAGGACAGACTGAGTGTGCTTTGGCCATTTAGGGGGCGC 1461  
QY 96 Ile-----GlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 110  
Db 1462 TTAGTAATGCCAAATACAGGGGT-GAGCTTACTATACACACGGTGCCATCCC 1514

## RESULT 50

US-09-969-532-5  
; Sequence 5, Application US/09969532  
; Patent No. 6777232  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides Encodin  
; FILE REFERENCE: LEX-0244-USA  
; CURRENT APPLICATION NUMBER: US/09/969,532  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US 60/237,280  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 1692  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-969-532-5

## Alignment Scores:

Pred. No.: 2.28 Length: 1692  
Score: 75.50 Matches: 28  
Percent Similarity: 41.84% Conservative: 13  
Best Local Similarity: 28.57% Mismatches: 40  
Query Match: 11.82% Indels: 18  
DB: 4 Gaps: 5

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-969-532-5 (1-1692)

QY 30 PheSerPro---GluAspLeuThrVallysValGlnAspPheVal----- 44  
Db 1255 TTTAACCCCTTTGTCGGACATCAAGTGAAGTCCAGAGCTCGTTTCATGGTTTCCCTGGGA 1314  
QY 45 -----GluIleHisGlyLysHisAsnGluArgGlnAspPheHisGly--- 58  
Db 1315 GTGCTGAGAGAGCTGAGTACACGGCAGAGAAATCATTCAGGACTTTTCCCATGGAAAC 1374  
QY 59 ---TyrIleSerArgGluPheHisArgTyrArgLeuPro-----SerAsnValAsp 75  
Db 1375 AACACAGCTTTAGTACAAATGCATCCAGAAATAAATGCCTTACATCCAAATCTGTCA 1434  
QY 76 GlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLys 95  
Db 1435 TCACCTCCCAACAGGACAGACTGAGGCAACTGGTGTCTTTGGCCATTTAGGGGGCGC 1494  
QY 96 Ile-----GlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 110  
Db 1495 TTAGTAATGCCAAATACAGGGGT-GAGCTTACTATACACACGGTGCCATCCC 1547

Search completed: May 30, 2005, 08:54:27  
Job time : 171.922 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: May 30, 2005, 05:14:41 ; Search time 864.74 Seconds  
(without alignments)  
873.287 Million cell updates/sec

Title: US-10-657-740-1\_COPY\_51\_173  
Perfect score: 639  
Sequence: 1 SLFRVLDSGISEVRSRDK.....HAERAIPVSRREKPTSAFSS 123

Scoring table:  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5700845 seqs, 3069779757 residues

Total number of hits satisfying chosen parameters: 11401690

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10657740/runat 27052005 165254 3380/app query.fasta\_1.590  
-DB=Published Applications NA -OPMT=fastap -SUFFIX=rnpb -MNMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNIT8=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=150 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10657740 @CGN 1 904 @runat 27052005 165254 3380  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	639	100.0	372	18	US-10-657-740-2	Sequence 2, Appli
2	599	93.7	1056	16	US-10-316-253-112	Sequence 112, App
3	577.5	90.4	1271	16	US-10-316-253-110	Sequence 110, App
4	567	88.7	531	13	US-10-105-427-1	Sequence 1, Appli
5	367	57.4	211	16	US-10-029-386-26581	Sequence 26581, A
6	367	57.4	573	16	US-10-029-386-12881	Sequence 12881, A
7	361	56.5	528	17	US-10-152-319A-1574	Sequence 1574, Ap
8	361	56.5	689	19	US-10-486-706-455	Sequence 455, App
9	361	56.5	1247	9	US-09-917-800A-1419	Sequence 1419, Ap
10	361	56.5	1247	19	US-10-486-706-105	Sequence 105, App
11	358	56.0	548	15	US-10-101-510-250	Sequence 250, App
12	354	55.4	528	18	US-10-770-668-35	Sequence 35, Appli
13	354	55.4	691	9	US-09-954-456-514	Sequence 514, App
14	354	55.4	691	10	US-09-960-706-869	Sequence 869, App
15	354	55.4	691	10	US-09-873-319-566	Sequence 566, App
16	354	55.4	691	10	US-09-873-367C-764	Sequence 764, App
17	354	55.4	691	17	US-10-133-937-61	Sequence 61, Appli
18	354	55.4	691	17	US-10-172-118-686	Sequence 686, App
19	354	55.4	691	17	US-10-159-563-61	Sequence 61, Appli
20	354	55.4	691	17	US-10-342-887-686	Sequence 686, App
21	354	55.4	691	19	US-10-843-641A-764	Sequence 764, App
22	354	55.4	691	19	US-10-843-641A-3541	Sequence 3541, App
23	354	55.4	691	19	US-10-367-057-187	Sequence 187, App
24	332.5	52.0	380	9	US-09-960-352-4143	Sequence 4143, Ap
25	283	44.3	449	9	US-09-960-352-4277	Sequence 4277, Ap
26	282.5	44.2	373	17	US-10-242-535A-26398	Sequence 26398, A
27	282.5	44.2	373	17	US-10-085-783A-26398	Sequence 26398, A
28	263	41.2	349	9	US-09-990-747-28	Sequence 28, Appli
29	261.5	40.9	764	18	US-10-605-498-91	Sequence 91, Appli
30	261.5	40.9	789	17	US-10-641-643-1300	Sequence 1300, App
31	261.5	40.9	847	14	US-10-153-668-283	Sequence 283, App
32	261.5	40.9	865	11	US-09-969-034-4480	Sequence 4480, Ap
33	261.5	40.9	865	17	US-10-172-118-626	Sequence 626, App
34	261.5	40.9	865	17	US-10-342-887-626	Sequence 626, App
35	261.5	40.9	1231	9	US-09-880-107-3865	Sequence 3865, Ap
36	261.5	40.9	1380	14	US-10-100-957A-169	Sequence 169, App
37	261	40.8	604	9	US-09-917-800A-981	Sequence 981, App
38	260	40.7	599	14	US-10-060-036-43	Sequence 43, Appli
39	260	40.7	634	19	US-10-784-420-43	Sequence 43, Appli
40	260	40.7	787	9	US-09-917-800A-1428	Sequence 1428, Ap
41	260	40.7	787	17	US-10-191-803-228	Sequence 228, App
42	260	40.7	787	17	US-10-152-319A-1963	Sequence 1963, Ap
43	259	40.5	396	9	US-09-960-352-12622	Sequence 12622, A
44	250.5	39.2	393	18	US-10-425-115-63403	Sequence 63403, A
45	250.5	39.2	480	14	US-10-226-956-320	Sequence 320, App
46	250.5	39.2	486	14	US-10-226-956-303	Sequence 303, App
47	250.5	39.2	1310	17	US-10-152-319A-1412	Sequence 1412, App
48	250.5	39.2	1820	18	US-10-723-860-5075	Sequence 5075, Ap
49	250	39.1	450	10	US-09-918-995-4949	Sequence 4949, Ap
50	248	38.8	503	10	US-09-911-504-146	Sequence 146, App
51	242	37.9	543	17	US-10-621-901-2279	Sequence 2279, Ap
52	241.5	37.8	708	10	US-09-814-353-14770	Sequence 14770, A
53	240.5	37.6	461	14	US-10-198-846-10360	Sequence 10360, A
54	240	37.6	563	14	US-10-198-846-8376	Sequence 8376, Ap
55	234.5	36.7	614	10	US-09-814-353-2040	Sequence 2040, Ap
56	234.5	36.7	614	10	US-09-814-353-8386	Sequence 8386, Ap
57	233.5	36.5	954	10	US-09-814-353-8386	Sequence 2073, A
58	229	35.8	1700	9	US-09-813-358-2	Sequence 2, Appli
59	229	35.8	1700	10	US-09-997-279-2	Sequence 2, Appli
60	226	35.4	421	10	US-09-918-995-4772	Sequence 4772, Ap
61	223.5	35.0	500	10	US-09-918-995-22134	Sequence 22134, A
62	222	34.7	400	10	US-09-918-995-4075	Sequence 4075, Ap
63	217	34.0	605	19	US-10-840-038-2	Sequence 2, Appli
64	217	34.0	736	19	US-10-840-038-3	Sequence 3, Appli
65	212	33.2	49634	11	US-09-968-007A-514	Sequence 514, App
66	212	33.2	49634	19	US-10-843-641A-6984	Sequence 6984, Ap
67	210.5	32.9	829	17	US-10-062-674-1924	Sequence 1924, Ap
68	209.5	32.8	874	17	US-10-133-937-62	Sequence 62, Appli
69	209.5	32.8	874	17	US-10-159-563-62	Sequence 62, Appli

70	32.4	387	18	US-10-723-860-380	Sequence 380, App
71	202	212	18	US-10-482-158-43	Sequence 43, Appl
72	199	381	18	US-10-425-115-173881	Sequence 173881, App
c 73	198.5	396	9	US-09-954-531-845	Sequence 845, App
c 74	198.5	396	9	US-09-954-531-1265	Sequence 1265, App
c 75	198.5	396	19	US-10-843-641A-1912	Sequence 1912, App
c 76	198.5	396	19	US-10-843-641A-2332	Sequence 2332, App
c 77	197.5	396	19	US-09-991-936-341	Sequence 341, App
78	195	549	10	US-10-060-036-34	Sequence 34, Appl
79	192.5	30.1	482	US-10-918-995-4433	Sequence 4433, App
80	192.5	30.1	550	US-10-621-901-2194	Sequence 2194, App
81	192	30.0	6806	US-10-205-331-105	Sequence 105, App
82	188	29.4	4206	US-10-101-510-288	Sequence 288, App
83	188	29.4	4206	US-10-283-975A-93	Sequence 93, Appl
84	182.5	28.6	410	US-10-468-091-35	Sequence 35, Appl
c 85	179	28.0	565	US-10-198-846-366	Sequence 366, App
c 86	174	27.2	510	US-10-363-345A-653	Sequence 653, App
c 87	174	27.2	510	US-10-363-345A-654	Sequence 654, App
c 88	174	27.2	510	US-10-363-483A-653	Sequence 654, App
c 89	174	27.2	510	US-10-363-483A-654	Sequence 654, App
c 90	172	26.9	542	US-10-621-901-2146	Sequence 2146, App
91	170	26.6	654	US-10-204-921-16	Sequence 16, Appl
92	165.5	25.9	487	US-10-369-493-30110	Sequence 30110, A
93	165.5	25.9	487	US-10-369-493-30111	Sequence 30111, A
94	163	25.5	445	US-09-918-995-10509	Sequence 10509, A
95	160	25.0	1495	US-10-006-285-269	Sequence 269, App
96	159.5	25.0	349	US-10-062-674-254	Sequence 254, App
97	159	24.9	1627	US-10-116-802-322	Sequence 322, App
98	159	24.9	2010	US-10-172-118-1482	Sequence 1482, App
99	159	24.9	2036	US-10-342-887-1482	Sequence 447, App
100	159	24.9	2036	US-10-006-285-447	Sequence 4925, App
101	158	24.7	810	US-10-723-860-4925	Sequence 3948, App
102	148	23.2	403	US-09-960-352-3948	Sequence 30109, A
103	148	23.2	490	US-10-369-493-30109	Sequence 30112, A
c 104	148	23.2	490	US-10-369-493-30112	Sequence 105, App
c 105	148	23.2	6806	US-10-205-331-105	Sequence 655, App
c 106	142.5	22.3	510	US-10-363-345A-655	Sequence 656, App
c 107	142.5	22.3	510	US-10-363-345A-656	Sequence 655, App
c 108	142.5	22.3	510	US-10-363-483A-655	Sequence 655, App
c 109	142.5	22.3	510	US-10-363-483A-656	Sequence 656, App
c 110	142	22.2	131	US-10-029-386-26899	Sequence 26899, A
c 111	142	22.2	451	US-10-102-524-483	Sequence 483, App
c 112	142	22.2	534	US-10-029-386-13199	Sequence 13199, A
113	137.5	21.5	450	US-09-880-107-3022	Sequence 3022, App
114	137.5	21.5	450	US-09-954-531-962	Sequence 962, App
115	137.5	21.5	450	US-09-954-531-1355	Sequence 1355, App
116	137.5	21.5	450	US-10-843-641A-2029	Sequence 2029, App
117	137.5	21.5	450	US-10-843-641A-2422	Sequence 2422, App
118	135.5	21.2	2632	US-10-044-090-312	Sequence 312, App
119	131	20.5	303	US-09-960-352-2010	Sequence 2010, App
c 120	131	20.5	395	US-09-918-995-4376	Sequence 4376, App
121	128	20.0	446	US-09-960-352-5348	Sequence 5348, App
122	127.5	20.0	2153	US-10-062-674-2106	Sequence 2106, App
123	125.5	19.6	1891	US-10-316-253-108	Sequence 108, App
124	124	19.2	339	US-09-867-701-6118	Sequence 6118, App
125	121	18.9	65	US-09-908-975-30348	Sequence 30348, A
c 126	120.5	18.9	42339	US-10-087-192-991	Sequence 991, App
127	118	18.5	442	US-09-918-995-5137	Sequence 5137, App
128	118	18.5	495	US-10-255-536-69	Sequence 69, Appl
129	115.5	18.1	478	US-09-918-995-20604	Sequence 20604, A
130	115	18.0	1922	US-10-104-047-1686	Sequence 1686, App
c 131	115	18.0	50002	US-10-087-192-994	Sequence 994, App
132	111.5	17.4	468	US-09-938-842A-2558	Sequence 2558, App
133	111.5	17.4	468	US-09-938-842A-2558	Sequence 2558, App
134	110.5	17.3	719	US-10-037-270-669	Sequence 669, App
135	110.5	17.3	719	US-10-117-722-669	Sequence 669, App
136	110.5	17.3	2152	US-10-723-860-395	Sequence 395, App
137	110.5	17.3	2410	US-10-104-047-972	Sequence 972, App
138	110.5	17.3	3503	US-10-723-860-5087	Sequence 5087, App
c 139	102.5	16.0	583	US-10-021-323-6441	Sequence 6441, App
c 140	102.5	16.0	597	US-10-021-323-5926	Sequence 5926, App
141	102.5	16.0	813	US-10-767-795-5171	Sequence 5171, App
142	101.5	15.9	943	US-10-310-154-324	Sequence 324, App

143	101.5	15.9	943	19	US-10-732-923-265	Sequence 265, App
144	101.5	15.9	1018	17	US-10-424-599-103855	Sequence 103855, A
c 145	101	15.8	60	10	US-09-908-975-15945	Sequence 15945, A
c 146	101	15.8	2731748	18	US-10-297-465A-1	Sequence 1, Appli
c 147	100.5	15.7	473	10	US-09-918-995-5122	Sequence 5122, Ap
c 148	99	15.5	768	17	US-10-425-114-9506	Sequence 9506, Ap
c 149	99	15.5	800	17	US-10-424-599-18307	Sequence 18307, A
150	98.5	15.4	599	18	US-10-021-323-5328	Sequence 5328, Ap

ALIGNMENTS

RESULT 1

US-10-657-740-2

Sequence 2, Application US/10657740

Publication NO. US20040157289A1

GENERAL INFORMATION:

APPLICANT: Salerno, John C.

APPLICANT: Hanna, Michael

APPLICANT: Koretz, Jane F.

APPLICANT: Crone, Donna

APPLICANT: Smith, Susan E.

TITLE OF INVENTION: PROTEIN EXPRESSION SYSTEM

FILE REFERENCE: 01794100406US1

CURRENT APPLICATION NUMBER: US/10/657,740

CURRENT FILING DATE: 2003-09-08

PRIOR FILING DATE: 2003-09-08

PRIOR APPLICATION NUMBER: US 60/408,680

PRIOR FILING DATE: 2002-09-06

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 372

TYPE: DNA

ORGANISM: Homo sapiens

US-10-657-740-2

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

7,71e-84

639.00

100.00%

100.00%

100.00%

18

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

372

123

0

0

0

0

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-657-740-2 (1-372)

Qy

1

SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys

20

Db

1

TCCTCTTCGCGACCGTCTGGACTCCGACATCTCTGAGTTTCGATCCGACCGGACAG

60

Qy

21

PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValLysValGln

40

Db

61

TTCTGTCATCTTCTCGATGTGAAGCATCTTCTCCGAGGACCTCACCGTGAAGTGCAG

120

Qy

41

AspAspPheValIleHisGlyIleHisGlyIleHisGlyIleHisGlyIleHisGlyIle

60

Db

121

GACGACTTTGTGAGATCCACGGAAGCACACGAGCCGAGGACGACCGGTCACCT

180

Qy

61

SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer

80

Db

181

TCCCTGTGAGTTCCACGCGCGTACCGCTTCCGCTTCCACGCTGACGACGACGACG

240

Qy

81

CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu

100

Db

241

TGCTCCCTCTCTGCGATGGCATGTCGTGACCTTCTGTGCGCCCAAGATCCAGACTGG

300

Qy

101

AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla

120

Db

301

GATGCCACCCACCGCGAGCGACCATCCCGTGTCTCGCGGAGGAGAGCCACCTCGCT

360

Qy

121

ProSerSer

123

Db

361

CCCTGCTCC

369

## RESULT 2

US-10-316-253-112  
 ; Sequence 112, Application US/10316253  
 ; Publication No. US20030162706A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Procter & Gamble Company  
 ; APPLICANT: Peters, Kevin  
 ; APPLICANT: Thompson, Larry  
 ; APPLICANT: Wang, Feng  
 ; APPLICANT: Greis, Kenneth  
 ; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
 ; FILE REFERENCE: 8865M  
 ; CURRENT APPLICATION NUMBER: US/10/316,253  
 ; CURRENT FILING DATE: 2002-12-10  
 ; PRIOR APPLICATION NUMBER: US 60/355,295  
 ; PRIOR FILING DATE: 2002-02-08  
 ; NUMBER OF SEQ ID NOS: 308  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 112  
 ; LENGTH: 1056  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (13)..(534)  
 ; OTHER INFORMATION:  
 ; US-10-316-253-112

Alignment Scores:  
 Pred. No.: 2,678-77 Length: 1056  
 Score: 539.00 Matches: 115  
 Percent Similarity: 97.56% Conservative: 5  
 Best Local Similarity: 93.50% Mismatches: 3  
 Query Match: 93.74% Indels: 0  
 DB: 16 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-316-253-112 (1-1056)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 163 TCTCTCTCCGACAGTGTGGACTCCGGCATCTCTGAGTCCGATCTGACCGGACCAAG 222  
 QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
 Db 223 TTGTGTCATCTCTGGATGTGAAGCACTTCTCTGAGGACCTCACCGTGAAGGTACTG 282  
 QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60  
 Db 283 GAAGATTTCGTGGATCCATGGCAACACACAGAGAGCGAGATGACCATGGCTACATT 342  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 343 TCCCGTGATTTCCCGTCGTACCGTCTGCTTCCATGTGGACCATGTCGCCCTCTCC 402  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 Db 403 TGCTCTCTGTCGGATGGCATGCTGACCTTCTCTGGCCCCAAGTCCAGTCTGGCTTG 462  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 463 GATGTCGGCCACAGCGAGGGGCCATTCCCGTGTACGCGGAGGAGAGGCCAGCTCGGCA 522  
 QY 121 ProSerSer 123  
 Db 523 CCTCGTCC 531

## RESULT 3

US-10-316-253-110  
 ; Sequence 110, Application US/10316253  
 ; Publication No. US20030162706A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Procter & Gamble Company

; APPLICANT: Peters, Kevin  
 ; APPLICANT: Thompson, Larry  
 ; APPLICANT: Wang, Feng  
 ; APPLICANT: Greis, Kenneth  
 ; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
 ; FILE REFERENCE: 8865M  
 ; CURRENT APPLICATION NUMBER: US/10/316,253  
 ; CURRENT FILING DATE: 2002-12-10  
 ; PRIOR APPLICATION NUMBER: US 60/355,295  
 ; PRIOR FILING DATE: 2002-02-08  
 ; NUMBER OF SEQ ID NOS: 308  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 110  
 ; LENGTH: 1271  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (159)..(749)  
 ; OTHER INFORMATION:  
 ; US-10-316-253-110

Alignment Scores:  
 Pred. No.: 5,058-74 Length: 1271  
 Score: 577.50 Matches: 115  
 Percent Similarity: 82.19% Conservative: 5  
 Best Local Similarity: 78.77% Mismatches: 3  
 Query Match: 90.38% Indels: 23  
 DB: 16 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-316-253-110 (1-1271)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGlu-----ValArgSerAsp 17  
 Db 309 TCTCTCTCCGACAGTGTGGACTCCGGCATCTCTGAGTCATGACCCCATATGTGGTTT 368  
 QY 14 -----ValArgSerAsp 17  
 Db 369 GTAATGACCAACACCATGCTGGAAACCCCAAGAACACCCCGCAAGGTCCGATCTGAC 428  
 QY 18 ArgAspLysPheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrVal 37  
 Db 429 CGGCAAGTTTGTGTCATCTCTTGGATGTGAAGCACTTCTCTCTGAGGACCTTCCCGTG 488  
 QY 38 LysValGlnAspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHis 57  
 Db 489 AAGGTACTGGAAGATTTCGTGGAGATCCATGGCAACACACAGAGAGCGAGTGCACAT 548  
 QY 58 GlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSer 77  
 Db 549 GGCTACATTTCCCGTGAATTTCCACCGTCGTACCGTCTGCTTCCAATGTGGACCATGCC 608  
 QY 78 AlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGln 97  
 Db 609 GCGCTCTCTCTGCTCTCTGTCGGATGGCATGTCGACCTTCTCTGGCCCCAAGGTCCAG 668  
 QY 98 ThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysPro 117  
 Db 659 TCTGCTGGATGCTGGCCACAGCGAGGGGCCCATTCCTGTCGCGGAGGAGAGGCC 728  
 QY 118 ThrSerAlaProSerSer 123  
 Db 729 AGCTCGGCACCTCTCGTCC 746

## RESULT 4

US-10-105-427-1  
 ; Sequence 1, Application US/10105427  
 ; Publication No. US20020177192A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Council of Scientific and Industrial Research  
 ; TITLE OF INVENTION: Chimeric protein alpha BNA crystallin with extraordinarily high  
 ; TITLE OF INVENTION: chaperone-like activity and a method thereof  
 ; FILE REFERENCE: US- 649

```
; CURRENT APPLICATION NUMBER: US/10/105,427
; CURRENT FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence for chimeric alpha BNAC
US-10-105-427-1

Alignment Scores:
Pred. No.: 4.89e-73 Length: 531
Score: 567.00 Matches: 106
Percent Similarity: 97.41% Conservativeness: 7
Best Local Similarity: 91.38% Mismatches: 3
Query Match: 88.73% Indels: 0
DB: 13 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-10-105-427-1 (1-531)

QY 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 184 GACACTGCTCTCAGAGATCGCGCTGGAGAGACAGAGTTCTCTGTCAACCTGGATGTG 243

QY 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHis 47
Db 244 AAGCACTTCTCCCGAGGAGCTCACCGTGAAGTGCAGGACCACTTTGTGGAGATCCAC 303

QY 48 GlyLysHisGlnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArg 67
Db 304 GGAAGACACAAACGAGCGCCAGGACGACACCGCTACATTTCCGCTGAGTTCCACCGCGC 363

QY 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
Db 364 TACGCGCTCGCTCCACACGCGGACCAAGTGGCCCTCTCTGTGCTCCCTGTCTCCGATGC 423

QY 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 424 ATGCTGACCTTCTGTGGCCCAAGATCCAGACTGGCTGGATGCCACCCACCGCCAGGCA 483

QY 108 AlaIleProValSerArgGluLysProThrSerAlaProSerSer 123
Db 484 GCCATCCCCGTGCGCGGAGGAGAGACCCACCTCGGCTCCCTCGTCC 531

RESULT 5
US-10-029-386-26581
; Sequence 26581, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26581
; LENGTH: 211
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR21.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
; OTHER INFORMATION: EST HUMAN HIT: BF726399.1, EVALUE 1.00e-113
; OTHER INFORMATION: SWISSPROT HIT: P02489, EVALUE 3.00e-36
; OTHER INFORMATION: EST_HUMAN HIT: BF726399.1, EVALUE 1.00e-113
```

```
; OTHER INFORMATION: NT HIT: gill4780619, EVALUE 1.00e-115
US-10-029-386-26581

Alignment Scores:
Pred. No.: 3e-44 Length: 211
Score: 367.00 Matches: 70
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.43% Indels: 0
DB: 16 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-10-029-386-26581 (1-211)

QY 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 73
Db 1 CAGGACGACCAACGCGCTACATTTCCGCTGAGTTCCACCGCGCTACCGCTGCCGCTCCAAC 60

QY 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 93
Db 61 GTGGACCAAGTCCGCGCTCTCTGTGCTCCCTGTCTGCGATGGCATGCTGACCTTCTGTGGC 120

QY 94 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113
Db 121 CCCAAGATCCAGACTGGCTGGATGCCACCCACCGGAGCGGACCATCCCCGCTGTGCGG 180

QY 114 GluLysProThrSerAlaProSerSer 123
Db 181 GAGGAGAGAGCCACCTCGGCTCCCTCGTCC 210

RESULT 6
US-10-029-386-12881
; Sequence 12881, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 12881
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR21.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
; OTHER INFORMATION: NT HIT: AF026952.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: BF726856.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P02489, EVALUE 1.00e-32
US-10-029-386-12881

Alignment Scores:
Pred. No.: 1.3e-43 Length: 573
Score: 367.00 Matches: 70
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.43% Indels: 0
DB: 16 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-10-029-386-12881 (1-573)

QY 54 GlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuProSerAsn 73
Db 203 CAGGACGACCAACGCGCTACATTTCCGCTGAGTTCCACCGCGCTACCGCTGCCGCTCCAAC 262
```

QY 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 93  
 Db 263 GTGGACCACTGGCGCTCTCTGCTCCCTGCTGCGATGGCATGCTGACCTCTCTGTGGC 322  
 QY 94 ProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProValSerArg 113  
 Db 323 CCCAAGATCCAGACTGGCTGATGCCACCAAGCGGAGGAGCCATCCCGCTGTGCGG 382  
 QY 114 GluGluLysProThrSerAlaProSerSer 123  
 Db 383 GAGGAGAGAGCCACCTCGGCTCCCTCGTCC 412

## RESULT 7

US-10-152-319A-1574  
 ; Sequence 1574, Application US/10152319A  
 ; Publication No. US20040072160A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mendrick, Donna  
 ; APPLICANT: Porter, Mark  
 ; APPLICANT: Johnson, Kory  
 ; APPLICANT: Higgins, Brandon  
 ; APPLICANT: Castle, Arthur  
 ; APPLICANT: Elashoff, Michael  
 ; TITLE OF INVENTION: Molecular Toxicology Modeling  
 ; FILE REFERENCE: 44921-5089-US  
 ; CURRENT APPLICATION NUMBER: US/10/152,319A  
 ; PRIOR FILING DATE: 2002-05-22  
 ; PRIOR APPLICATION NUMBER: US 60/292,335  
 ; PRIOR FILING DATE: 2001-05-22  
 ; PRIOR APPLICATION NUMBER: US 60/297,523  
 ; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/298,925  
 ; PRIOR FILING DATE: 2001-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/303,810  
 ; PRIOR FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: US 60/303,807  
 ; PRIOR FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: US 60/303,808  
 ; PRIOR FILING DATE: 2001-07-10  
 ; PRIOR APPLICATION NUMBER: US 60/315,047  
 ; PRIOR FILING DATE: 2001-08-28  
 ; PRIOR APPLICATION NUMBER: US 60/324,928  
 ; PRIOR FILING DATE: 2001-09-27  
 ; PRIOR APPLICATION NUMBER: US 60/330,867  
 ; PRIOR FILING DATE: 2001-11-01  
 ; PRIOR APPLICATION NUMBER: US 60/330,462  
 ; PRIOR FILING DATE: 2001-10-22  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 2221  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 1574  
 ; LENGTH: 528  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. NM\_012935  
 US-10-152-319A-1574

Alignment Scores:  
 Pred. No.: 8,76e-43 Length: 528  
 Score: 361.00 Matches: 67  
 Percent Similarity: 79.49% Conservative: 26  
 Best Local Similarity: 57.26% Mismatches: 18  
 Query Match: 56.49% Indels: 6  
 DB: 17 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-152-319A-1574 (1-528)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 181 ATTGACACTGGGCTCTCAGAGATGGTATGGAGAGGACAGGTTCTCTGTGAACCTGGAC 240  
 QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIle 46

Db 241 GTGAGACACTTCTCTCCAGAGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAGGTG 300  
 QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
 Db 301 CACGGCAAGCAGAGAGCGCCAGCAACATCGCTTCATCTCCAGGGAGTTCACACAG 360  
 QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 361 AAGTACCGGATCCAGCGGAGCGTGGATCTCTCACCATTTACTTCTCCCTGTGCATCGAT 420  
 QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 421 GGAGTCTCTCACTGTGAATGGCAAGGAACAG-----GCCTCTGGCCCTGAG 468  
 QY 107 ArgAlaIleProValSerArgGluGluLysPro-----ThrSerAlaPro 121  
 Db 469 CGCACCATTCCTCCATCACCCGCTGAAGAGAGAGCCCTGTCTCCTGCAGGCCCT 519

## RESULT 8

US-10-486-706-455  
 ; Sequence 455, Application US/10486706  
 ; Publication No. US20050071088A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LANDFIELD, PHILIP W.  
 ; APPLICANT: BLALOCK, ERIC M.  
 ; APPLICANT: CHEN, KUBY-CHU  
 ; APPLICANT: FOSTER, THOMAS C.  
 ; TITLE OF INVENTION: GENE EXPRESSION PROFILE BIOMARKERS AND THERAPEUTIC TARGETS FOR  
 ; TITLE OF INVENTION: BRAIN AGING AND AGE-RELATED COGNITIVE IMPAIRMENT  
 ; FILE REFERENCE: 50229-426  
 ; CURRENT APPLICATION NUMBER: US/10/486,706  
 ; CURRENT FILING DATE: 2004-02-13  
 ; PRIOR APPLICATION NUMBER: PCT/US02/25607  
 ; PRIOR FILING DATE: 2002-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/311,343  
 ; PRIOR FILING DATE: 2001-08-13  
 ; NUMBER OF SEQ ID NOS: 461  
 ; SOFTWARE: Patent in version 3.2  
 ; SEQ ID NO 455  
 ; LENGTH: 689  
 ; TYPE: DNA  
 ; ORGANISM: Rattus rattus  
 US-10-486-706-455

## Alignment Scores:

Pred. No.: 1.3e-42 Length: 689  
 Score: 361.00 Matches: 67  
 Percent Similarity: 79.49% Conservative: 26  
 Best Local Similarity: 57.26% Mismatches: 18  
 Query Match: 56.49% Indels: 6  
 DB: 19 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-486-706-455 (1-689)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 193 ATTGACACTGGGCTCTCAGAGATGGTATGGAGAGGACAGGTTCTCTGTGAACCTGGAC 252  
 QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIle 46  
 Db 253 GTGAGACACTTCTCTCCAGAGAACTCAAAGTCAAGGTTCTGGGAGACGTGATTGAGGTG 312  
 QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
 Db 313 CACGGCAAGCAGAGAGCGCCAGCAACATCGCTTCATCTCCAGGGAGTTCACACAG 372  
 QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 373 AAGTACCGGATCCAGCGGAGCGTGGATCTCTCACCATTTACTTCTCCCTGTGCATCGAT 432  
 QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106





```
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 250
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-10-101-510-250

Alignment Scores:
Pred. No.: 2,558-42 Length: 548
Score: 358.00 Matches: 67
Percent Similarity: 78.63% Conservative: 25
Best Local Similarity: 57.26% Mismatches: 19
Query Match: 56.03% Indels: 6
DB: 15 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x US-10-101-510-250 (1-548)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26
DB 201 ATTGACCTGGACTCTCAGAGATGCGCTGGAGAGGACAGGTTCTGTCAACCTGGAT 260
QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLe 46
DB 261 GTGAAGCACTTCTCCCGAGAGAGCTCAAGGTCAAAGTGTGGTGTGATGATTGAGGTG 320
QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66
DB 321 CACGCGCAACATGAAGAGCGCCAGATGAACATGTTTCATCTCCAGGAGTTCCACAGG 380
QY 67 ArgTyrArgLeuProSerAsnValArgGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86
DB 381 AAATACCGGATCCAGCTGATGTGACCTCTCACCATTACTTCATCTCCTGTCTGAT 440
QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106
DB 441 GGGGTCTCTCACTGTGAATGGACCAAGGAAGCAA-----GCCCTGGCCCGAG 488
QY 107 ArgAlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121
DB 489 CGCACCATTCCTCAATACCGGTGAAGAGAAGCTGTGTCTCACTGCAGCCCC 539

RESULT 12
US-10-770-668-35
; Sequence 35, Application US/10770668
; Publication No. US20040191843A1
; GENERAL INFORMATION:
; APPLICANT: Wright, Susan C.
; APPLICANT: Larrick, James W.
; APPLICANT: Nock, Steffen R.
; APPLICANT: Wilson, David S.
; TITLE OF INVENTION: Cell-Killing Molecules and Methods of Use Thereof
; FILE REFERENCE: ABSALUS-08602
; CURRENT APPLICATION NUMBER: US/10/770,668
; CURRENT FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 35
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-770-668-35

; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 250
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-10-101-510-250

Alignment Scores:
Pred. No.: 2,558-42 Length: 548
Score: 358.00 Matches: 67
Percent Similarity: 78.63% Conservative: 25
Best Local Similarity: 57.26% Mismatches: 19
Query Match: 56.03% Indels: 6
DB: 15 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x US-10-101-510-250 (1-548)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26
DB 201 ATTGACCTGGACTCTCAGAGATGCGCTGGAGAGGACAGGTTCTGTCAACCTGGAT 260
QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluLe 46
DB 261 GTGAAGCACTTCTCCCGAGAGAGCTCAAGGTCAAAGTGTGGTGTGATGATTGAGGTG 320
QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66
DB 321 CACGCGCAACATGAAGAGCGCCAGATGAACATGTTTCATCTCCAGGAGTTCCACAGG 380
QY 67 ArgTyrArgLeuProSerAsnValArgGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86
DB 381 AAATACCGGATCCAGCTGATGTGACCTCTCACCATTACTTCATCTCCTGTCTGAT 440
QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106
DB 441 GGGGTCTCTCACTGTGAATGGACCAAGGAAGCAA-----GCCCTGGCCCGAG 488
QY 107 ArgAlaIleProValSerArgGluLysPro-----ThrSerAlaPro 121
DB 489 CGCACCATTCCTCAATACCGGTGAAGAGAAGCTGTGTCTCACTGCAGCCCC 539

RESULT 13
US-09-954-456-514
; Sequence 514, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 514
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-514
```



```
Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 449 GTCTCTCACTGTAATGACCAAGAAACAG-----GTCTTGGCCCTGAGCGC 496

Qy 108 AlaIleProValSerArgGluGluysPro-----ThrSerAlaPro 121
Db 497 ACCATTCCCATCACCCCGTGAAGAGAGCCTGTGTGTACCGCAGCCCC 544

RESULT 16
US-10-657-740-1_COPY_51_173 (1-123) x US-09-873-367C-764 (1-691)
; Sequence 764, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: Signature Gene Sets
; CURRENT APPLICATION NUMBER: US/09/873,367C
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 764
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-764

Alignment Scores:
Pred. No.: 1,39e-41 Length: 691
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservative: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 10 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x US-09-873-367C-764 (1-691)
Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspHisPheValIlePheLeuAspVal 27
Db 209 GACACTGGACTCTCAGAGATGGCTGGAGAGACAGAGTTCTGTCTCAACTGGATGTG 268

Qy 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIleHis 47
Db 269 AAGCACTTCTCCCGAGAGAACTCAAAGTTAAGGTGTTGGAGATGTGATTGAGGTGCAT 328

Qy 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 67
Db 329 GGAATAACATGAAGAGCGCCAGATGAACATGTTTCATCTCCAGGAGTTCCACAGGAAA 388

Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
Db 389 TACCGATCCCGATGATGATAGACCTCTCACCATTACTTCTCCTCTGCTATCTGATGGG 448

Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 449 GTCTCTCACTGTAATGACCAAGAAACAG-----GTCTTGGCCCTGAGCGC 496

Qy 108 AlaIleProValSerArgGluGluysPro-----ThrSerAlaPro 121
Db 497 ACCATTCCCATCACCCCGTGAAGAGAGCCTGTGTGTACCGCAGCCCC 544

RESULT 17
US-10-657-740-1_COPY_51_173 (1-123) x US-10-133-937-61 (1-691)
; Sequence 61, Application US/10133937
; Publication No. US20030207278A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
; TITLE OF INVENTION: OTHER BIOLOGICAL STATES
; FILE REFERENCE: 11613.56US01
; CURRENT APPLICATION NUMBER: US/10/133,937
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-937-61

Alignment Scores:
Pred. No.: 1,39e-41 Length: 691
Score: 354.00 Matches: 66
Percent Similarity: 78.45% Conservative: 25
Best Local Similarity: 56.90% Mismatches: 19
Query Match: 55.40% Indels: 6
DB: 17 Gaps: 2

US-10-657-740-1_COPY_51_173 (1-123) x US-10-133-937-61 (1-691)
Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspHisPheValIlePheLeuAspVal 27
Db 209 GACACTGGACTCTCAGAGATGGCTGGAGAGACAGAGTTCTGTCTCAACTGGATGTG 268

Qy 28 LysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIleHis 47
Db 269 AAGCACTTCTCCCGAGAGAACTCAAAGTTAAGGTGTTGGAGATGTGATTGAGGTGCAT 328

Qy 48 GlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 67
Db 329 GGAATAACATGAAGAGCGCCAGATGAACATGTTTCATCTCCAGGAGTTCCACAGGAAA 388

Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGly 87
Db 389 TACCGATCCCGATGATGATAGACCTCTCACCATTACTTCTCCTCTGCTATCTGATGGG 448

Qy 88 MetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArg 107
Db 449 GTCTCTCACTGTAATGACCAAGAAACAG-----GTCTTGGCCCTGAGCGC 496

Qy 108 AlaIleProValSerArgGluGluysPro-----ThrSerAlaPro 121
Db 497 ACCATTCCCATCACCCCGTGAAGAGAGCCTGTGTGTACCGCAGCCCC 544

RESULT 18
US-10-172-118-686
; Sequence 686, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mac, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
```









; APPLICANT: Bvatt, John C.  
 ; APPLICANT: Mathialagan, Nagesan  
 ; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 ; FILE REFERENCE: 16511.006/37-21(10298)C  
 ; CURRENT APPLICATION NUMBER: US/09/960,352  
 ; CURRENT FILING DATE: 2001-09-24  
 ; NUMBER OF SEQ ID NOS: 15112  
 ; SEQ ID NO 4277  
 ; LENGTH: 449  
 ; TYPE: DNA  
 ; ORGANISM: Bos taurus  
 ; OTHER INFORMATION: Clone ID: 19-BOVMS1-021-Q1-E1-E3  
 US-09-960-352-4277

Alignment Scores:  
 Pred. No.: 1.98e-31 Length: 449  
 Score: 283.00 Matches: 48  
 Percent Similarity: 86.25% Conservativeness: 21  
 Best Local Similarity: 60.00% Mismatches: 11  
 Query Match: 44.29% Indels: 0  
 DB: 9 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-960-352-4277 (1-449)

QY 7 LeuAppSerGlyIleSerGluValArgSerAspArgAspPheValIlePheLeuApp 26  
 DB 210 ATTGACACTTCTCCAGAGATCGCTCTGGAGAGACAGATTTCTCTCAACCTGGAT 269  
 QY 27 ValIleHisPheSerProGluAspLeuThrValIleValGlnAspPheValGluIle 46  
 DB 270 GTGAGACACTTCTCCAGAGATCGCTCTGGAGATCGCTCTGGAGATCGCTCTGGAG 329  
 QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyIleSerArgGluPheHisArg 66  
 DB 330 CATGCAACATGAAGAGCGCCAGATGAACATGTTTATCTCCGGAGTTCCACAGG 389  
 QY 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 DB 390 AAATACCGGATCCAGCTGACGTGGAACCTCTCGCCATTACTTCTCCCTGCTGCTGAT 449

RESULT 26

US-10-242-535A-26398  
 ; Sequence 26398, Application US/0242535A  
 ; Publication No. US20040013663A1  
 ; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.  
 ; APPLICANT: Liew, C.C.  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
 ; FILE REFERENCE: 4231/2005  
 ; CURRENT APPLICATION NUMBER: US/10/242,535A  
 ; PRIOR FILING DATE: 2002-09-12  
 ; PRIOR APPLICATION NUMBER: US 10/085,783  
 ; PRIOR FILING DATE: 2002-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/305,340  
 ; PRIOR FILING DATE: 2001-07-13  
 ; PRIOR APPLICATION NUMBER: US 60/275,017  
 ; PRIOR FILING DATE: 2001-03-12  
 ; PRIOR APPLICATION NUMBER: US 60/271,955  
 ; PRIOR FILING DATE: 2001-02-28  
 ; NUMBER OF SEQ ID NOS: 58994  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 26398  
 ; LENGTH: 373  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (24)..(24)  
 ; OTHER INFORMATION: n is a, c, g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (227)..(227)  
 ; OTHER INFORMATION: n is a, c, g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (227)..(227)

; OTHER INFORMATION: n is a, c, g, or t  
 US-10-242-535A-26398

Alignment Scores:  
 Pred. No.: 1.78e-31 Length: 373  
 Score: 282.50 Matches: 60  
 Percent Similarity: 71.30% Conservativeness: 22  
 Best Local Similarity: 52.17% Mismatches: 19  
 Query Match: 44.21% Indels: 15  
 DB: 17 Gaps: 4

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-242-535A-26398 (1-373)

QY 14 ValArgSerAspArgAspPheValIlePheLeuAspValIleHisPheSerProGlu 33  
 DB 3 ATGCGCTGGAGAGACAGGNTCTCTCAACCTGGATGTGAAGCACTTCTCCCGAGAG 62  
 QY 34 AspLeuThrValIleValGlnAspPheValGluIleHisGlyLysHisGluArg 53  
 DB 63 GAACCAAGATTAAAGTTGGAGATGTGGATGTGATGATGATGATGATGATGATGATGAT 122  
 QY 54 GlnAspHisGlyIleSerArgGluPheHisArgTyArgLeuProSerAsn 73  
 DB 123 CAGGATGAACATGTTTCTCTCCAGGAGTTCCACAGGAATACCGGATCCAGCTGAT 182  
 QY 74 ValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPheCysGly 93  
 DB 183 GTAGACCTCTCCACATTACTTCTCCCTGCTCATCTGATGGG-----GGN 227  
 QY 94 ProLysIle-----GlnThrGlyLeuAspAlaThrHisAlaGluArgAla 108  
 DB 228 CCTCACTGTGAATGGACCAAGGAAACAGGTCTCTG-GCCCTGAGCCCA-----CCA 280  
 QY 109 IleProValSerArgGluLysPro-----ThrSerAlaPro 121  
 DB 281 TTTCCTCATCCCGTGAAGAGAGCTGTGTCTCAGCGAGCCCC 325

RESULT 27

US-10-085-783A-26398  
 ; Sequence 26398, Application US/10085783A  
 ; Publication No. US20040037841A1  
 ; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.  
 ; APPLICANT: Liew, C.C.  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
 ; FILE REFERENCE: 4231/2002  
 ; CURRENT APPLICATION NUMBER: US/10/085,783A  
 ; CURRENT FILING DATE: 2002-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/305,340  
 ; PRIOR FILING DATE: 2001-07-13  
 ; PRIOR APPLICATION NUMBER: US 60/275,017  
 ; PRIOR FILING DATE: 2001-03-12  
 ; PRIOR APPLICATION NUMBER: US 60/271,955  
 ; PRIOR FILING DATE: 2001-02-28  
 ; NUMBER OF SEQ ID NOS: 58994  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 26398  
 ; LENGTH: 373  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (24)..(24)  
 ; OTHER INFORMATION: n is a, c, g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (227)..(227)  
 ; OTHER INFORMATION: n is a, c, g, or t  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (227)..(227)

US-10-085-783A-26398

Alignment Scores:  
 Pred. No.: 1.78e-31 Length: 373  
 Score: 282.50 Matches: 60



;; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL  
;; GENE EXPRESSION

;; NUMBER OF SEQUENCES: 1508

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

;; STREET: 3174 PORTER DRIVE

;; CITY: PALO ALTO

;; STATE: CALIFORNIA

;; COUNTRY: USA

;; ZIP: 94304

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/10/641,643

;; FILING DATE: 14-AUG-2003

;; CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: <Unknown>

;; FILING DATE: <Unknown>

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Zeller, Karen J.

;; REGISTRATION NUMBER: 37,071

;; REFERENCE/DOCKET NUMBER: PA-0001 US

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (650) 855-0555

;; TELEFAX: (650) 845-4166

;; INFORMATION FOR SEQ ID NO: 1300:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 789 base pairs

;; TYPE: nucleic acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; IMMEDIATE SOURCE:

;; LIBRARY: GENBANK

;; CLONE: G32477

;; SEQUENCE DESCRIPTION: SEQ ID NO: 1300 :

US-10-641-643-1300

Alignment Scores:

Pred. No.: 6,51e-28 Length: 789  
Score: 261.50 Matches: 55  
Percent Similarity: 65.18% Conservative: 18  
Best Local Similarity: 49.11% Mismatches: 30  
Query Match: 40.92% Indels: 10  
DB: 17 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-641-643-1300 (1-789)

Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
Db 293 CTCACGACGGGGTCTCGAGATCGGCACACTGCGGACCGGTGCGCGGTGTCCTGGAT 352  
Qy 27 VallysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 46  
Db 353 GTCAACCACTTCGCCCGGACGAGCTGACGGTCAAGACCAAGATGGCGTGGGAGATC 412  
Qy 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
Db 413 ACCGCAAGCAGCAGGAGCGGCGAGCAGCATGCTCATCTCCCGGTGCTTACGCGG 472  
Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 473 AAATACACGCTGCCCGCGGTGTGAGACCCCAAGTTTCCTCTCCTGTGCTTCCCTGAG 532  
Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
Db 533 GGCACTGACCGTGA-GGCCCC-----CATGCCCAA 564  
Qy 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
Db 107 ArgAlaIleProValSerArgGluGluLysProThr 118

Db 565 GCTAGCCAGCGAGTCCAAACGAGATCACCATCCCACT 600

RESULT 31

US-10-153-668-283

;; Sequence 283, Application US/10153668

;; Publication No. US20030092616A1

;; GENERAL INFORMATION:

;; APPLICANT: HONDA, Goichi

;; APPLICANT: MATSUDA, Akio

;; APPLICANT: MURAMATSU, Shuji

;; APPLICANT: ISHIZAWA, Kenya

;; TITLE OF INVENTION: STAT6 Activating Gene

;; FILE REFERENCE: 1254-0207P

;; CURRENT APPLICATION NUMBER: US/10/153,668

;; PRIOR FILING DATE: 2002-05-24

;; PRIOR FILING DATE: 2001-05-25

;; PRIOR APPLICATION NUMBER: US 60/316,031

;; PRIOR FILING DATE: 2001-08-31

;; PRIOR APPLICATION NUMBER: US 60/328,403

;; PRIOR FILING DATE: 2001-10-12

;; PRIOR APPLICATION NUMBER: JP 2001-157043

;; PRIOR FILING DATE: 2001-05-25

;; PRIOR APPLICATION NUMBER: JP 2001-260681

;; PRIOR FILING DATE: 2001-08-30

;; PRIOR APPLICATION NUMBER: JP 2001-313175

;; PRIOR FILING DATE: 2001-10-10

;; NUMBER OF SEQ ID NOS: 488

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 283

;; LENGTH: 847

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: CDS

;; LOCATION: (108)..(722)

US-10-153-668-283

Alignment Scores:

Pred. No.: 7,22e-28 Length: 847  
Score: 261.50 Matches: 55  
Percent Similarity: 65.18% Conservative: 18  
Best Local Similarity: 49.11% Mismatches: 30  
Query Match: 40.92% Indels: 10  
DB: 14 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-153-668-283 (1-847)

Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
Db 348 CTCACGACGGGGTCTCGAGATCGGCACACTGCGGACCGGTGCGCGGTGTCCTGGAT 407  
Qy 27 VallysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGlu 46  
Db 408 GTCAACCACTTCGCCCGGACGAGCTGACGGTCAAGACCAAGATGGCGTGGGAGATC 467  
Qy 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
Db 468 ACCGCAAGCAGCAGGAGCGGCGAGCAGCATGCTCATCTCCCGGTGCTTACGCGG 527  
Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 528 AAATACACGCTGCCCGCGGTGTGAGACCCCAAGTTTCCTCTCCTGTGCTTCCCTGAG 587  
Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
Db 588 GGCACTGACCGTGA-GGCCCC-----CATGCCCAA 619  
Qy 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
Db 620 GCTAGCCAGCGAGTCCAAACGAGATCACCATCCCACT 655

RESULT 32

```

US-09-969-034-4480
; Sequence 4480, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catano, Theodore J.
; APPLICANT: Dwivedi, Poonima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4480
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-034-4480

Alignment Scores:
Pred. No.: 7,45e-28 Length: 865
Score: 261.50 Matches: 55
Percent Similarity: 65.18% Conservative: 18
Best Local Similarity: 49.11% Mismatches: 30
Query Match: 40.92% Indels: 10
DB: 11 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-09-969-034-4480 (1-865)
Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26
Db 348 CTCAGCAGCGGGTCTCGGAGATCCGGACACTGCGGACCGTGGCGGTGCTCCCTGGAT 407
Qy 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIle 46
Db 408 GTCAACCACTTCGCCCGGACGAGCTGACGGTCAAGACCAAGATGGCGTGGTGAGATC 467
Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 66
Db 468 ACCGGAAGCAGCAGGAGCGGACGAGCATGCTACATCTCCCTCCCTGCTCCCTGAG 527
Qy 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86
Db 528 AAATACACGCTGCCCGGACGAGCGGACGAGCATGCTACATCTCCCTCCCTGCTCCCTGAG 587
Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106
Db 588 GGCACACTGACCGTGA-GGCCCC-----CATGCCCA 619

RESULT 33
US-10-172-118-626
; Sequence 626, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 626
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_001540
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-626

Alignment Scores:
Pred. No.: 7,45e-28 Length: 865
Score: 261.50 Matches: 55
Percent Similarity: 65.18% Conservative: 18
Best Local Similarity: 49.11% Mismatches: 30
Query Match: 40.92% Indels: 10
DB: 11 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-10-172-118-626 (1-865)
Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26
Db 348 CTCAGCAGCGGGTCTCGGAGATCCGGACACTGCGGACCGTGGCGGTGCTCCCTGGAT 407
Qy 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIle 46
Db 408 GTCAACCACTTCGCCCGGACGAGCTGACGGTCAAGACCAAGATGGCGTGGTGAGATC 467
Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 66
Db 468 ACCGGAAGCAGCAGGAGCGGACGAGCATGCTACATCTCCCTCCCTGCTCCCTGAG 527
Qy 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86
Db 528 AAATACACGCTGCCCGGACGAGCGGACGAGCATGCTACATCTCCCTCCCTGCTCCCTGAG 587
Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106
Db 588 GGCACACTGACCGTGA-GGCCCC-----CATGCCCA 619

RESULT 34
US-10-342-887-626
; Sequence 626, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118

```

```

; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 626
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-626

Alignment Scores:
Pred. No.: 7,45e-28 Length: 865
Score: 261.50 Matches: 55
Percent Similarity: 65.18% Conservative: 18
Best Local Similarity: 49.11% Mismatches: 30
Query Match: 40.92% Indels: 10
DB: 17 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-10-342-887-626 (1-865)

Qy 7 LeuaspSerGlyIleSerProGluValArgSerAspArgAspLysPheValIlePheLeuasp 26
Db 348 CTCAGACGCGGGTCTCGAGATCCGGCACACTGCGGACCGCTGCGCGTGTCCCTGGAT 407

Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspAspPheValGluile 46
Db 408 GTCAACCACTTCCCGCGAGCGAGCTGACGTCAAGACCAAGATGCGGTGTGGAGATC 467

Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrlleSerArgGluPheHisArg 66
Db 468 ACCGGCAAGCAGCAGGAGCGGCGAGCAGCAGTGTGACACCCCAAGTTTCCTCCTCCTGTCCTGGAT 527

Qy 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaasp 86
Db 528 AAATACACGCTGCCCGCGGCTGAGACCCCAAGTTTCCTCCTCCTGTCCTGGAT 587

Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAlaPheHisArg 106
Db 588 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 619

Qy 107 ArgAlaIleProValSerArgGluGluLysProthr 118
Db 620 GCTAGCCACGCGAGTCCACGAGATCACCATCCCACT 655

RESULT 35
US-09-880-107-3865
; Sequence 3865, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3865
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 Z23090
US-09-880-107-3865

Alignment Scores:
Pred. No.: 1.25e-27 Length: 1231
Score: 261.50 Matches: 55
Percent Similarity: 65.18% Conservative: 18

US-10-657-740-1_COPY_51_173 (1-123) x US-10-342-887-626 (1-865)

Qy 7 LeuaspSerGlyIleSerProGluValArgSerAspArgAspLysPheValIlePheLeuasp 26
Db 1003 CTCAGACGCGGGTCTCGAGATCCGGCACACTGCGGACCGCTGCGCGTGTCCCTGGAT 1062

Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspAspPheValGluile 46
Db 1063 GTCAACCACTTCCCGCGAGCGAGCTGACGTCAAGACCAAGATGCGGTGTGGAGATC 1122

Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrlleSerArgGluPheHisArg 66

```

```

Best Local Similarity: 49.11% Mismatches: 30
Query Match: 40.92% Indels: 10
DB: 9 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-09-880-107-3865 (1-1231)

Qy 7 LeuaspSerGlyIleSerProGluValArgSerAspArgAspLysPheValIlePheLeuasp 26
Db 732 CTCAGACGCGGGTCTCGAGATCCGGCACACTGCGGACCGCTGCGCGTGTCCCTGGAT 791

Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspAspPheValGluile 46
Db 792 GTCAACCACTTCCCGCGAGCGAGCTGACGTCAAGACCAAGATGCGGTGTGGAGATC 851

Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrlleSerArgGluPheHisArg 66
Db 852 ACCGGCAAGCAGCAGGAGCGGCGAGCAGCAGTGTGACACCCCAAGTTTCCTCCTCCTGTCCTGGAT 911

Qy 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaasp 86
Db 912 AAATACACGCTGCCCGCGGCTGAGACCCCAAGTTTCCTCCTCCTGTCCTGGAT 971

Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAlaPheHisArg 106
Db 972 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 1003

Qy 107 ArgAlaIleProValSerArgGluGluLysProthr 118
Db 1004 GCTAGCCACGCGAGTCCACGAGATCACCATCCCACT 1039

RESULT 36
US-10-100-957A-169
; Sequence 169, Application US/10100957A
; Publication No. US20030096322A1
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1A
; CURRENT APPLICATION NUMBER: US/10/100,957A
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 169
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GFP-HSP27
; NAME/KEY: CDS
; LOCATION: (1)..(1380)
US-10-100-957A-169

Alignment Scores:
Pred. No.: 1.48e-27 Length: 1380
Score: 261.50 Matches: 55
Percent Similarity: 65.18% Conservative: 18
Best Local Similarity: 49.11% Mismatches: 30
Query Match: 40.92% Indels: 10
DB: 14 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-10-100-957A-169 (1-1380)

Qy 7 LeuaspSerGlyIleSerProGluValArgSerAspArgAspLysPheValIlePheLeuasp 26
Db 1003 CTCAGACGCGGGTCTCGAGATCCGGCACACTGCGGACCGCTGCGCGTGTCCCTGGAT 1062

Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspAspPheValGluile 46
Db 1063 GTCAACCACTTCCCGCGAGCGAGCTGACGTCAAGACCAAGATGCGGTGTGGAGATC 1122

Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrlleSerArgGluPheHisArg 66

```

Db 1123 ACCGGCAAGCAGGAGCGGAGGAGGATGCTACATCTCCGGTGTTCACCGG 1182  
 QY 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 1183 AAATACAGCTGCCCGGCGGTGTGGACCCCAAGTTTCTCTCTCCCTGTCCCTGAG 1242  
 QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 1243 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 1274  
 QY 107 ArgAlaIleProValSerArgGluGluLysProThr 118  
 Db 1275 GCTAGCCAGCGAGTCCCAACGAGATCACCATCCCACT 1310

## RESULT 37

US-09-917-800A-981/C  
 ; Sequence 981, Application US/09917800A  
 ; Patent No. US20020119462A1  
 ; GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna  
 ; APPLICANT: Porter, Mark  
 ; APPLICANT: Johnson, Kory  
 ; APPLICANT: Castle, Arthur  
 ; APPLICANT: Elashoff, Michael  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Molecular Toxicology Modeling  
 ; FILE REFERENCE: 44921-5038-US  
 ; CURRENT APPLICATION NUMBER: US/09/917,800A  
 ; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/222,040  
 ; PRIOR FILING DATE: 2000-07-31  
 ; PRIOR APPLICATION NUMBER: US 60/222,880  
 ; PRIOR FILING DATE: 2000-11-02  
 ; PRIOR APPLICATION NUMBER: US 60/290,029  
 ; PRIOR FILING DATE: 2001-05-11  
 ; PRIOR APPLICATION NUMBER: US 60/290,645  
 ; PRIOR FILING DATE: 2001-05-15  
 ; PRIOR APPLICATION NUMBER: US 60/292,336  
 ; PRIOR FILING DATE: 2001-05-22  
 ; PRIOR APPLICATION NUMBER: US 60/295,798  
 ; PRIOR FILING DATE: 2001-06-06  
 ; PRIOR APPLICATION NUMBER: US 60/297,457  
 ; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/298,884  
 ; PRIOR FILING DATE: 2001-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/303,459  
 ; PRIOR FILING DATE: 2001-07-09  
 ; NUMBER OF SEQ ID NOS: 1740  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 981  
 ; LENGTH: 604  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020119462A1 A1176658  
 US-09-917-800A-981

## Alignment Scores:

Pred. No.: 5,2e-28 Length: 604  
 Score: 261.00 Matches: 57  
 Percent Similarity: 65.18% Conservative: 16  
 Best Local Similarity: 50.89% Mismatches: 33  
 Query Match: 40.85% Indels: 6  
 DB: 9 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-917-800A-981 (1-604)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 492 CTCAGTACGGGTGCTCAGAGATCCGACAGCGCCGATCTCGCCCGGTGCTCCCTGGAC 433  
 QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIle 46

Db 432 GTCAACCACTTCGCTCCTGAGGAGCTCACAGTTAAGACCAAGGAGCGGTGTGGAGATC 373  
 QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrlleSerArgGluPheHisArg 66  
 Db 372 ACTGCGAAGCAGCAAGAAAGGAGGATGAACATGGCTACATCTCTCGGTGCTTCACCGG 313  
 QY 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 312 AAATACAGCTCCCTCCAGGTGTGGACCCCACTTGGTGTCTCTTCCCTGTCCCTGAG 253  
 QY 87 GlyMetLeuThrPheCysGly-----ProLysIleGlnThrGlyLeuAspAlaThrHis 104  
 Db 252 GGCACACTCAGCGTGAAGGCTCCGTCGCCAAAGCAGTCACA-----CAATCA 205  
 QY 105 AlaGluArgAlaIleProValSerArgGluGluLys 116  
 Db 204 GCGGAGATCACCATTCCGTCACCTTTCAGGCCCGGT 169

## RESULT 38

US-10-060-036-43  
 ; Sequence 43, Application US/10060036  
 ; Publication No. US20030073144A1  
 ; GENERAL INFORMATION:

; APPLICANT: Benson, Darin R.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Persing, David H.  
 ; APPLICANT: Hepler, William T.  
 ; APPLICANT: Jjiang, Yugu  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; FILE REFERENCE: 210121.566  
 ; CURRENT APPLICATION NUMBER: US/10/060,036  
 ; CURRENT FILING DATE: 2002-01-30  
 ; NUMBER OF SEQ ID NOS: 4560  
 ; SOFTWARE: Fast-Seq for Windows Version 4.0  
 ; SEQ ID NO 43  
 ; LENGTH: 599  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-060-036-43

## Alignment Scores:

Pred. No.: 7,21e-28 Length: 599  
 Score: 260.00 Matches: 55  
 Percent Similarity: 65.18% Conservative: 18  
 Best Local Similarity: 49.11% Mismatches: 30  
 Query Match: 40.69% Indels: 10  
 DB: 14 Gaps: 1

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-060-036-43 (1-599)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
 Db 254 CTCAGCAGCGGGTCTCGAGATCCGCGACACCTCGGACCGGTGCGCGGTGCTCCCTGGAT 313  
 QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspPheValGluIle 46  
 Db 314 GTCAACCACTTCGCCCCGAGCGAGCTGAGCGTCAAGACCAAGATGGCGTGTGGAGATC 373  
 QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrlleSerArgGluPheHisArg 66  
 Db 374 ACCGGCAAGCAGCAGGAGCGGCGAGCAGCATGCTACATCTCCCGGTGCTTCACCGG 433  
 QY 67 ArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
 Db 434 AAATACAGCTGCCCGGCGGTGTGGACCCCACTTTCCTCTCCCTGTCCCTGAG 493  
 QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106  
 Db 494 GGCACACTGACCGTGA-GGCCCC-----CATGCCCAA 525

```
Qy 107 ArgAlaIleProValSerArgGluGluLysProThr 118
Db 526 GCTAGCCACGCGTCTCAACGAGATCACCATCCCACT 561
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1428
; LENGTH: 787
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 M86389
US-09-917-800A-1428
Alignment Scores:
Pred. No.: 1,08e-27 Length: 787
Score: 260.00 Matches: 57
Percent Similarity: 65.18% Conservative: 16
Best Local Similarity: 50.89% Mismatches: 33
Query Match: 40.69% Indels: 6
DB: 9 Gaps: 2
US-10-657-740-1_COPY_51_173 (1-123) x US-09-917-800A-1428 (1-787)
Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26
Db 296 CTCAGCAGCGGTCTCGAGATCCGACAGACGCGCGGTGTCCTGCGGTGTCCTCGAC 355
Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspAspPheValGluIle 46
Db 356 GTCACCACTTCGCTCTCGAGAGCTCACATGTTAAGCAAGGCGGTGTCGAGATC 415
Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyIleSerArgGluPheHisArg 66
Db 416 ACTGGCAAGCAGCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 475
Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86
Db 476 AATACACGCTCCCTCCAGGTGTGACCCACCCCTGCTCTCTCTCTCTCTCTCTCTCT 535
Qy 87 GlyMetLeuThrPheCysGly-----ProLysIleGlnThrGlyLeuAspAlaThrHis 104
Db 536 GGCACACTCACCGTGGAGGCTCCGCTGCCCAAGCAGTCACA-----CAATCA 583
Qy 105 AlaGluArgAlaIleProValSerArgGluLys 116
Db 584 GCGGAGATCACCAATTCGGTCACTTTTCGAGGCCCGT 619
RESULT 41
US-10-191-803-228
; Sequence 228, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
Qy 107 ArgAlaIleProValSerArgGluGluLysProThr 118
Db 526 GCTAGCCACGCGTCTCAACGAGATCACCATCCCACT 561
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1428
; LENGTH: 787
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-764-420-43
Alignment Scores:
Pred. No.: 7,84e-28 Length: 634
Score: 260.00 Matches: 57
Percent Similarity: 65.18% Conservative: 16
Best Local Similarity: 50.89% Mismatches: 33
Query Match: 40.69% Indels: 6
DB: 19 Gaps: 2
US-10-657-740-1_COPY_51_173 (1-123) x US-10-764-420-43 (1-634)
Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26
Db 250 CTCAGCAGCGGTCTCGAGATCCGACAGACGCGGTGTCCTGCGGTGTCCTCGAC 309
Qy 27 VallysHisPheSerProGluAspLeuThrVallysValGlnAspAspPheValGluIle 46
Db 310 GTCAACCACTTCGCTCCGAGAGCTCACATGTAAGCAAGGAGGCGGTGTCGAGATC 369
Qy 47 HisGlyLysHisAsnGluArgGlnAspAspHisGlyIleSerArgGluPheHisArg 66
Db 370 ACTGCAAGCAGCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 429
Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86
Db 430 AATACACGCTCCCTCCAGGTGTGACCCACCCCTAGTGTCTCTCTCTCTCTCTCTCT 489
Qy 87 GlyMetLeuThrPheCysGly-----ProLysIleGlnThrGlyLeuAspAlaThrHis 104
Db 490 GGCACACTTCGCTGGAGGCTCCGTTGCCCAAGCAGTCACG-----CAGTCA 537
Qy 105 AlaGluArgAlaIleProValSerArgGluLys 116
Db 538 GCGGAGATCACCAATTCGGTCACTTTTCGAGGCCCGT 573
RESULT 40
US-09-917-800A-1428
; Sequence 1428, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Caslie, Arthur
; APPLICANT: Elashoff, Michael
```



APPLICANT: PORTER, Mark  
APPLICANT: JOHNSON, Kory  
APPLICANT: HIGGS, Brandon  
APPLICANT: CASTLE, Arthur  
APPLICANT: ELASHOFF, Michael  
TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling  
FILE REFERENCE: 44921-5090US  
CURRENT APPLICATION NUMBER: US/10/191,803  
CURRENT FILING DATE: 2002-07-10  
PRIOR APPLICATION NUMBER: US 60/303,819  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: US 60/305,623  
PRIOR FILING DATE: 2001-07-17  
PRIOR APPLICATION NUMBER: US 60/369,351  
PRIOR FILING DATE: 2002-04-03  
PRIOR APPLICATION NUMBER: US 60/377,611  
PRIOR FILING DATE: 2002-05-06  
NUMBER OF SEQ ID NOS: 1140  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 228  
LENGTH: 787  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM\_031970  
US-10-191-803-228

Alignment Scores:  
Pred. No.: 1.08e-27 Length: 787  
Score: 260.00 Matches: 57  
Percent Similarity: 65.18% Conservative: 16  
Best Local Similarity: 50.89% Mismatches: 33  
Query Match: 40.69% Indels: 6  
DB: 17 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-191-803-228 (1-787)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
Db 314 CTCAGACGGGTGTGTGAGAGATCCGACAGCGCGGATCGTGGCGGTGCTCCCTGGAC 373  
QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIle 46  
Db 374 GTCAACCACTTCCTCTCAGAGCTCAGAGTGAAGACCAAGAGGCGGTGGAGATC 433  
QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
Db 434 ACTGGCAAGCAGCAAGAAAGGAGGATGACATGCTATCTCTCGGTGCTTCACCGG 493  
QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 494 AAATACACGCTCCCTCCAGGTGTGACCCACCTTGGTGTCTCTCTCCCTGCTCCCTGAG 553  
QY 87 GlyMetLeuThrPheCysGly-----ProLysIleGlnThrGlyLeuAspAlaThrHis 104  
Db 554 GGCACACTCAGCGTGGAGGCTCCGCTGCCAAAGCAGTCACA-----CAATCA 601  
QY 105 AlaGluArgAlaIleProValSerArgGluLys 116  
Db 602 GCGGAGATCACCATTCCGGTCACITTCGAGGCCCGT 637

RESULT 42  
US-10-152-319A-1963  
Sequence 1963, Application US/10152319A  
Publication No. US20040072160A1  
GENERAL INFORMATION:

APPLICANT: Mendrick, Donna  
APPLICANT: Porter, Mark  
APPLICANT: Johnson, Kory  
APPLICANT: Higgs, Brandon  
APPLICANT: Castle, Arthur  
APPLICANT: Elashoff, Michael  
TITLE OF INVENTION: Molecular Toxicology Modeling

FILE REFERENCE: 44921-5089-US  
CURRENT APPLICATION NUMBER: US/10/152,319A  
CURRENT FILING DATE: 2002-05-22  
PRIOR APPLICATION NUMBER: US 60/292,335  
PRIOR FILING DATE: 2001-05-22  
PRIOR APPLICATION NUMBER: US 60/297,523  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/298,925  
PRIOR FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/303,810  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: US 60/303,807  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: US 60/303,808  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: US 60/315,047  
PRIOR FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: US 60/324,928  
PRIOR FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: US 60/330,867  
PRIOR FILING DATE: 2001-11-01  
PRIOR APPLICATION NUMBER: US 60/330,462  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 2221  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1963  
LENGTH: 787  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. NM\_031970  
US-10-152-319A-1963

Alignment Scores:

Pred. No.: 1.08e-27 Length: 787  
Score: 260.00 Matches: 57  
Percent Similarity: 65.18% Conservative: 16  
Best Local Similarity: 50.89% Mismatches: 33  
Query Match: 40.69% Indels: 6  
DB: 17 Gaps: 2

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-10-152-319A-1963 (1-787)

QY 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26  
Db 314 CTCAGACGGGTGTGTGAGAGATCCGACAGCGCGGATCGTGGCGGTGCTCCCTGGAC 373  
QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIle 46  
Db 374 GTCAACCACTTCCTCTCAGAGCTCAGAGTGAAGACCAAGAGGCGGTGGAGATC 433  
QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66  
Db 434 ACTGGCAAGCAGCAAGAAAGGAGGATGACATGCTATCTCTCGGTGCTTCACCGG 493  
QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86  
Db 494 AAATACACGCTCCCTCCAGGTGTGACCCACCTTGGTGTCTCTCTCCCTGCTCCCTGAG 553  
QY 87 GlyMetLeuThrPheCysGly-----ProLysIleGlnThrGlyLeuAspAlaThrHis 104  
Db 554 GGCACACTCAGCGTGGAGGCTCCGCTGCCAAAGCAGTCACA-----CAATCA 601  
QY 105 AlaGluArgAlaIleProValSerArgGluLys 116  
Db 602 GCGGAGATCACCATTCCGGTCACITTCGAGGCCCGT 637

RESULT 43

US-09-960-352-12622  
Sequence 12622, Application US/09960352  
Patent No. US20020137139A1  
GENERAL INFORMATION:

```

US-10-657-740-1_COPY_51_173 (1-123) x US-10-425-115-63403 (1-393)
Qy 7 LeuAspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAsp 26
Db 9 CTCAGCAGCGGGTCTCGAGATCGGCACACTCGGACCGCTGGCGGCTGCCCTCGGAT 68
Qy 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIle 46
Db 69 GTCAACCACTTCGCCCGCGACCGAGCTGACGGTCAAGACCAAGGATGGCGTGTGAGATC 128
Qy 47 HisGlyLysHisAsnGluValArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArg 66
Db 129 ACCGGCTAGCACGAGGACGCGACGACGATGGCTACATCTCCCGGTGCTTCACGGCG 188
Qy 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86
Db 189 AAATACACGCTGCCCGCGGTGTGGACCCACCCNAGTTTCTCTCTCTCTCTCTCTCTG 248
Qy 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106
Db 249 GGCACACTGACCGTGA-GGCCCC-----CATGCCCA 280
Qy 107 ArgAlaIleProValSerArgGluGluLysProThr 118
Db 281 GCTAGCCACGCGAGTCCAACGAGATCACCATCCCACT 316

RESULT 45
US-10-226-956-320
; Sequence 320, Application US/10226956
; Publication No. US20030060399A1
; GENERAL INFORMATION:
; APPLICANT: Brophy, Colleen
; APPLICANT: Komalavilas, Padmini
; APPLICANT: Panitch, Alyssa
; APPLICANT: Joshi, Lokesh
; APPLICANT: Seal, Brandon L.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
; FILE REFERENCE: ASU-1061-US
; CURRENT APPLICATION NUMBER: US/10/226,956
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/314,535
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 320
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence encoding human HSP20
US-10-226-956-320

Alignment Scores:
Pred. No.: 1.29e-26 Length: 480
Score: 250.50 Matches: 46
Percent Similarity: 63.64% Conservative: 24
Best Local Similarity: 41.82% Mismatches: 25
Query Match: 39.20% Indels: 15
DB: 14 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-10-226-956-320 (1-480)
Qy 11 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPhe 30
Db 190 -GTGCTCAGGTACCAACTGACCCACCGCCACTTCTCCGTTTATTAGACGTGAACACTTT 249
Qy 31 SerProGluAspLeuThrValLysValGlnAspAspPheValGluIleHisGlyLysHis 50
Db 250 AGCCCAAGAGATGACGAGTCNAAGTTGTAGGAGAGCATGTGGAAAGTTTCACCGAGACAT 309
Qy 51 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeu 70

```

```
Db 310 GAAGAGACACGAGTGAACATGTTTCGTAGCGAGAGAAATCCATCGACGGTATCGTCG 369
QY 71 ProSerAnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 90
Db 370 CCCCAGAGTCGATCTCGACGCTGTGACGAGTCATTTACCGCTGAGGAGTGCTCAGT 429
QY 91 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 110
Db 430 ATC-----CAAGCAGGCCCC 444
QY 111 ValSerArgGluGluLysProThrSerAla 120
Db 445 CGGTACGCCCAAGCCCGCTTCGGCTGCT 474

RESULT 46
US-10-226-956-303
; Sequence 303, Application US/10226956
; Publication No. US20030060399A1
; GENERAL INFORMATION:
; APPLICANT: Brophy, Colleen
; APPLICANT: Komalavilas, Padmini
; APPLICANT: Panitch, Alyssa
; APPLICANT: Joshi, Lokesh
; APPLICANT: Seal, Brandon L.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
; FILE REFERENCE: ASU-1061-US
; CURRENT APPLICATION NUMBER: US/10/226,956
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/314,535
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 320
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 303
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence encoding rat HSP20
US-10-226-956-303

Alignment Scores:
Pred. No.: 1,32e-26 Length: 486
Score: 250.50 Matches: 48
Percent Similarity: 63.39% Conservative: 23
Best Local Similarity: 42.86% Mismatches: 26
Query Match: 39.20% Indels: 15
DB: 14 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-10-226-956-303 (1-486)
QY 12 SerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSer 31
Db 193 GCCCAAGTCCCGACCGATCCAGGCTATTTCAGCGTCTGTAGACGTAAGCATTTAGT 252
QY 32 ProGluAspLeuThrValLysValGlnAspPheValGluIleHisGlyLysHisAsn 51
Db 253 CCAGAGAAATTCAGTAAAGTAGTGGAGACCATGTCGAGGTACATGTCAGACGAA 312
QY 52 GluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 71
Db 313 GAGACCTGTATGAACACGGTTTCATCGTCGAGAGTTTCACCGCGCTTATCGCTTCCG 372
QY 72 SerAnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 91
Db 373 CCGGGGTGTGATCCCGCGCGCTGCATCATCATCATCATCATCATCATCATCATCATCAT 432
QY 92 CysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProVal 111
Db 433 -----CAAGCCACCGGCC 447
QY 112 SerArgGluGluLysProThrSerAlaProSerSer 123
Db 448 TCTGCTCAGGCTCGCTTCCATCCCTCTCCGCGCA 483
```

```
RESULT 47
US-10-152-319A-1412
; Sequence 1412, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1412
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. D29960
US-10-152-319A-1412
```

```
Alignment Scores:
Pred. No.: 5,65e-26 Length: 1310
Score: 250.50 Matches: 48
Percent Similarity: 63.39% Conservative: 23
Best Local Similarity: 42.86% Mismatches: 26
Query Match: 39.20% Indels: 15
DB: 17 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-10-152-319A-1412 (1-1310)
QY 12 SerGluValArgSerAspArgAspLysPheValIlePheLeuAspValLysHisPheSer 31
Db 198 GCCCAGGTCCCGACCGATCCAGGCTATTTCGTGCTGCTGATGTGAAGCATTTCTCG 257
QY 32 ProGluAspLeuThrValLysValGlnAspPheValGluIleHisGlyLysHisAsn 51
Db 258 CCAGAGAAATCTCTGTCAAGCTGTTGTTGTCATGCCATGTGGAGTCCATGCTCGGATGAG 317
QY 52 GluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgArgTyrArgLeuPro 71
Db 318 GAGCCCGCAGATGACATGGATTCATGCTCAGAGTTCCACCGCCGATACCGCTGCT 377
QY 72 SerAnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThrPhe 91
Db 378 CTTGGCGTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
```

```
Qy 92 CysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIleProVal 111
Db 438 -----CAGGCCACACCGCG 452
Qy 112 SerArgGluGluLysProThrSerAlaProSerSer 123
Db 453 TCGGCCAGGCTCACTTCGCTCACCACCTGCTGCC 488

RESULT 48
US-10-723-860-5075
; Sequence 5075, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; TITLE OF INVENTION: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5075
; LENGTH: 1820
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-5075

Alignment Scores:
Pred. No.: 9,16e-26 Length: 1820
Score: 250.50 Matches: 46
Percent Similarity: 63.64% Conservative: 24
Best Local Similarity: 41.82% Mismatches: 25
Query Match: 39.20% Indels: 15
DB: 18 Gaps: 1

US-10-657-740-1_COPY_51_173 (1-123) x US-10-723-860-5075 (1-1820)
Qy 11 IleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspValIleHisPhe 30
Db 212 GTCGCCAGGTCGCCAGCGACCGCCGCGCCTTCCTGGTGTCTAGACGTGAAGCACTTC 271
Qy 31 SerProGluAspLeuThrValIleValGlnAspPheValGluIleHisGlyLysHis 50
Db 272 TCGCCGAGGAAATTCGTCTAGGTGTGGCGGACACGTGGAGGTGCACGCGCGGCAC 331
Qy 51 AsnGluArgGlnAspAspHisGlyTyrIleSerArgGluPheHisArgGlyArgLeu 70
Db 332 GAGGAGCGCCCGATGACACGGATTCTCGCGCGCGAGTTCACCGCTGCTACCGCCTG 391
Qy 71 ProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAspGlyMetLeuThr 90
Db 392 CCGCTGGCTGGATTCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 451
Qy 91 PheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGluArgAlaIlePro 110
Db 452 ATC-----CAGGCCGCGCA 466

RESULT 49
US-09-918-995-4949
; Sequence 4949, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
```

```
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4949
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(450)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-4949

Alignment Scores:
Pred. No.: 1.39e-26 Length: 450
Score: 250.00 Matches: 44
Percent Similarity: 80.52% Conservative: 18
Best Local Similarity: 57.14% Mismatches: 15
Query Match: 39.12% Indels: 0
DB: 10 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x US-09-918-995-4949 (1-450)
Qy 8 AspSerGlyIleSerGluValArgSerAspArgAspLysPheValIlePheLeuAspVal 27
Db 220 GACACTGGACTCTCAAGATCGCGCTGGAGAGACAGGGTCTCTGTCAACTGGATGTG 279
Qy 28 LysHisPheSerProGluAspLeuThrValIleValGlnAspPheValGluIleHis 47
Db 280 AAGCACTTCTCCCGAGAGAACTCAGAGTTAATGTGTGTGATGATGAGATTGAGGTGCAT 339
Qy 48 GlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 67
Db 340 GGAATAACATGAAGAGCGCGAGGATGAACATGGTTTCATCTCCAGGGAGTTCACAGGACA 399
Qy 68 TyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSer 84
Db 400 TACCGGATCCGCGTGTGTAGACCCCTGTGACCATTTACTTCACTCCCTGTCTCN 450

RESULT 50
US-09-911-904-146
; Sequence 146, Application US/09911904
; Publication No. US20030096234A1
; GENERAL INFORMATION:
; APPLICANT: Farr, Spencer B.
; APPLICANT: Pickett, Gavin G.
; APPLICANT: Neft, Robin Eileen
; APPLICANT: Dunn, II, Robert Thomas
; TITLE OF INVENTION: CANINE TOXICITY GENES
; FILE REFERENCE: 400742000200
; CURRENT APPLICATION NUMBER: US/09/911,904
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/220,057
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-911-904-146

Alignment Scores:
Pred. No.: 3.23e-26 Length: 503
Score: 248.00 Matches: 53
Percent Similarity: 65.42% Conservative: 17
Best Local Similarity: 49.53% Mismatches: 28
Query Match: 38.81% Indels: 10
DB: 10 Gaps: 1
```

US-10-657-740-1\_COPY\_51\_173 (1-123) x US-09-911-904-146 (1-503)

```
QY 7 LeuAspSerGlyIleSerGluValArgSerAspLysPheValIlePheLeuAsp 26
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 CTCAGACGGCGGTCTCGAGATCCGCGACGCGCGCGTGGCGGTCTCCCTGGAC 265
QY 27 ValLysHisPheSerProGluAspLeuThrValLysValGlnAspAspPheValGluIle 46
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 GTCAACCACTTCGCCCCGAGGAGCTGACGGTCAAGACGAGGACGCGGTGGAGATA 325
QY 47 HisGlyLysHisAsnGluArgGlnAspHisGlyTyrIleSerArgGluPheHisArg 66
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
326 ACTGGCAAGCACGAAGAGAGGAGGATGAGCATGGTACATCTCCCGCCGCTCACTCCC 385
QY 67 ArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSerCysSerLeuSerAlaAsp 86
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
386 AATACACCCCTGCCCTGGTGTGGATCCTACCTGGTCTCTCTCTCTCTCTCTCTCTGAG 445
QY 87 GlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeuAspAlaThrHisAlaGlu 106
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
446 GGCACCTCTCAG-----GTGGAGGCTCC-CATGCCAA 477
QY 107 ArgAlaIleProValSerArg 113
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
478 GCCAGCCACCAGTCGGCAGA 498
```

Search completed: May 30, 2005, 09:27:43  
Job time : 873.74 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 30, 2005, 04:00:28 ; Search time 2317.47 Seconds  
(without alignments)  
2020.266 Million cell updates/sec

Title: US-10-657-740-1\_COPY\_51\_173

Perfect score: 639

Sequence: 1 SLFRTVLDSGISEVRSRDK.....HAERAIPVSRREKTSAPSS 123

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10657740/runat\_27052005\_165253\_3329/app\_query.fasta\_1.590  
-DB=EST -OPMT=fastcap -SUFFIX=first -MINMATCH=0.1 -IOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=150  
-DOALIGN=200 -THRM=score -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTPMT=ptc -NORM=ext -HAPSIX=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10657740@cgn\_1\_1\_5533 @runat\_27052005\_165253\_3329 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FCGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hc:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gse1:  
9: gb\_gse2:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	639	100.0	438	2	BF727355 by20d03.y
2	639	100.0	446	4	BM697673 UI-E-DXO-
3	639	100.0	460	2	BF727464 by21h03.y
4	639	100.0	466	2	BF727220 by18d12.y
5	639	100.0	500	2	BF727239 by18g02.y
6	639	100.0	508	4	BM696489 UI-E-DXO-
7	639	100.0	511	4	BM697090 UI-E-DXO-
8	639	100.0	512	4	BM697099 UI-E-DXO-
9	639	100.0	518	2	BF726854 by13a10.y

10	639	100.0	522	9	AY419529 Homo sapi
11	639	100.0	523	2	BF726890 by13f01.y
12	639	100.0	533	4	BM704974 UI-E-DXO-
13	639	100.0	534	4	BM705774 UI-E-DXO-
14	639	100.0	536	2	BF726399 by06a09.y
15	639	100.0	558	2	BF726679 by10d02.y
16	639	100.0	569	4	BM697368 UI-E-DXO-
17	639	100.0	576	2	BF726236 by03f01.y
18	639	100.0	577	2	BF727002 by15c11.y
19	639	100.0	578	4	BM721893 UI-E-E00-
20	639	100.0	592	2	BF727295 by19e10.y
21	639	100.0	592	4	BM705926 UI-E-DXO-
22	639	100.0	593	4	BM697101 UI-E-DXO-
23	639	100.0	596	4	BM696581 UI-E-DXO-
24	639	100.0	599	4	BM696976 UI-E-DXO-
25	639	100.0	607	6	CD675250 fs21c02.y
26	639	100.0	612	6	CD676130 fs32e02.y
27	639	100.0	629	2	BF727028 by15g05.y
28	639	100.0	630	4	BM696651 UI-E-DXO-
29	639	100.0	631	4	BM706251 UI-E-DXO-
30	639	100.0	633	6	CD672144 fg10b07.y
31	639	100.0	635	4	BM697066 UI-E-DXO-
32	639	100.0	659	4	BM706270 UI-E-DXO-
33	639	100.0	661	4	BM697160 UI-E-DXO-
34	639	100.0	681	5	EX118596 EX118596
35	639	100.0	698	2	BF727324 by19h12.y
36	639	100.0	724	4	BM722336 UI-E-E00-
37	635	99.4	559	4	BM722779 UI-E-E00-
38	633	99.1	532	2	BF726358 by05d12.y
39	632	98.9	522	9	AY419530 Pan trogl
40	631	98.7	519	4	BM696477 UI-E-DXO-
41	630	98.6	577	2	BF726253 by03h05.y
42	630	98.6	587	2	BF726422 by06d05.y
43	629	98.4	597	2	BF726330 by05b01.y
44	629	98.4	604	5	BQ640267 he26c10.y
45	626	98.0	580	4	BM706139 UI-E-DXO-
46	626	98.0	586	4	BM722650 UI-E-E00-
47	626	98.0	607	4	BM696670 UI-E-DXO-
48	620	97.0	512	4	BM706160 UI-E-DXO-
49	620	97.0	544	2	BF726363 by05e10.y
50	616	96.4	498	4	BM696853 UI-E-DXO-
51	616	96.4	523	2	BF726438 by06f01.y
52	616	96.4	588	4	BM696799 UI-E-DXO-
53	615	96.2	663	4	BM686748 UI-E-CQO-
54	611	95.6	371	4	BM705517 UI-E-DXO-
55	611	95.6	303	2	BF727205 by18b11.y
56	611	95.6	558	4	BM696667 UI-E-DXO-
57	601	94.1	688	7	CF732528 UI-M-HA0-
58	600	93.9	536	4	BM686206 UI-E-C11-
59	599	93.7	447	6	CB848427 M2PN-4253
60	599	93.7	486	7	CK627569 io05f11.y
61	599	93.7	522	9	AY419531 Mus muscu
62	599	93.7	542	5	BQ637175 he06f05.y
63	599	93.7	552	7	CK628252 ip09h08.y
64	599	93.7	556	6	CB848450 M2PN-4276
65	599	93.7	557	7	CK627503 io04g04.y
66	599	93.7	558	7	CK627631 io06f03.y
67	599	93.7	566	7	CK627600 io06c02.y
68	599	93.7	570	7	CK628219 ip09d07.y
69	599	93.7	574	7	CK627249 io01b06.y
70	599	93.7	579	7	CK628200 ip09b11.y
71	599	93.7	581	7	CO429470 UI-N-HW0-
72	599	93.7	582	7	CK628246 ip09g11.y
73	599	93.7	586	7	CK628391 ip11f11.y
74	599	93.7	587	7	CK627945 ip04h08.y
75	599	93.7	588	7	CK628525 ip13f04.y
76	599	93.7	590	7	CK627884 ip04a08.y
77	599	93.7	590	7	CK628336 ip10h08.y
78	599	93.7	590	7	CK628526 ip13f05.y
79	599	93.7	593	7	CK628271 ip10b05.y
80	599	93.7	594	7	CK628372 ip11d12.y
81	599	93.7	600	4	BG805475 0693-23 M
82	599	93.7	600	4	BG808732 2121-59 M

```

83 599 93.7 600 4 BI991006
84 599 93.7 600 6 CA527708
85 599 93.7 601 7 CK627483
86 599 93.7 601 7 CK627500
87 599 93.7 601 7 CK627510
88 599 93.7 602 7 CK627644
89 599 93.7 606 6 CB849766
90 599 93.7 606 6 CK628546
91 599 93.7 615 7 CK627740
92 599 93.7 617 6 CB848065
93 599 93.7 621 6 CB845999
94 599 93.7 622 7 CK627736
95 599 93.7 623 7 CK627777
96 599 93.7 626 7 CK627695
97 599 93.7 626 7 CK627737
98 599 93.7 629 6 CB055541
99 599 93.7 632 6 CB842313
100 599 93.7 633 7 CK627814
101 599 93.7 636 7 CK627616
102 599 93.7 636 7 CK627776
103 599 93.7 637 6 CB848191
104 599 93.7 639 7 CK627374
105 599 93.7 642 7 CK627338
106 599 93.7 643 6 CB846376
107 599 93.7 643 7 CK627445
108 599 93.7 643 7 CO429892
109 599 93.7 645 6 CB848071
110 599 93.7 645 7 CK627651
111 599 93.7 659 6 CB841917
112 599 93.7 662 6 CB841336
113 599 93.7 665 6 CB849833
114 599 93.7 669 6 CB840463
115 599 93.7 669 6 CB840784
116 599 93.7 676 6 CB840784
117 599 93.7 688 7 CO427248
118 599 93.7 690 6 CB846475
119 599 93.7 690 7 CN455827
120 599 93.7 690 7 CO428179
121 599 93.7 691 7 CF731811
122 599 93.7 692 6 CB845729
123 599 93.7 696 7 CO429233
124 599 93.7 699 7 CO427006
125 599 93.7 699 7 CO427334
126 599 93.7 702 6 CB844707
127 599 93.7 705 7 CO429072
128 599 93.7 705 7 CO429948
129 599 93.7 709 7 CO428729
130 599 93.7 721 7 CF733985
131 599 93.7 721 7 CF731519
132 599 93.7 760 7 CO427314
133 599 93.7 763 7 CO427704
134 599 93.7 771 7 CO427433
135 599 93.7 773 7 CO428121
136 599 93.7 798 7 CO427114
137 599 93.7 834 6 CB845009
138 599 93.7 851 7 CB845124
139 598 93.6 677 7 CN435619
140 598 93.6 745 7 CN441369
141 598 93.6 809 7 CN438019
142 598 93.6 836 7 CR550416
143 598 93.6 897 7 CR550671
144 597 93.4 618 4 BM697176
145 595 93.1 573 6 CB844502
146 595 93.1 698 7 CO429320
147 594 93.0 478 4 BM705752
148 594 93.0 854 6 CB848776
149 593 92.8 619 7 CK627238
150 593 92.8 625 7 CK627598

BI991006 4083-92 M
CA527708 8051-64 M
CK627483 1004d12.Y
CK627500 1004f12.Y
CO429510 UI-M-HW0-
CK627644 1006h01.Y
CB849766 MRA-1515
CK628546 1p13h12.Y
CK627740 1p02b06.Y
CB848065 M2PN-3875
CB845999 M2PN-1715
CK627736 1p02b02.Y
CK627777 1p02f06.Y
CK627695 1p01e09.Y
CK627737 1p02b03.Y
CB055541 NISC-JJ06
CB842313 M15E-2958
CK627814 1p03b02.Y
CK627616 1006d08.Y
CK627776 1p02f05.Y
CB848191 M2PN-4009
CK627374 1002h07.Y
CK627338 1002b04.Y
CB846376 M2PN-2110
CK627445 1003h09.Y
CO429892 UI-M-HW0-
CB848071 M2PN-3891
CK627651 1006h10.Y
CB841917 M15E-2546
CB841336 M15E-1855
CB849833 MRA-1582
CB840463 M15E-0760
CB840784 M15E-1296
CB840784 M15E-1183
CO427248 UI-M-HW0-
CB846475 M2PN-2213
CN455827 UI-M-HF0-
CO428179 UI-M-HW0-
CF731811 UI-M-HA0-
CB845729 M2PN-1430
CO429233 UI-M-HW0-
CO427006 UI-M-HW0-
CO427334 UI-M-HW0-
CB844707 M2PN-0259
CO429072 UI-M-HW0-
CO429948 UI-M-HW0-
CO428729 UI-M-HW0-
CF733985 UI-M-HA0-
CF731519 UI-M-HA0-
CO427314 UI-M-HW0-
CO427704 UI-M-HV0-
CO427433 UI-M-HW0-
CO428121 UI-M-HW0-
CO427114 UI-M-HW0-
CB845009 M2PN-0575
CB845124 M2PN-0694
CN435619 BE030013B
CN441369 BE04025B2
CN438019 BE04014A2
CB550416 CR550416
CB550671 CR550671
BM697176 UI-E-DW0-
CB844502 M2PN-0044
CO429320 UI-M-HW0-
BM705752 UI-E-DW0-
CB848776 MRA-0308
CK627238 1001a02.Y
CK627598 1006b10.Y

```

BF727355 438 bp mRNA linear EST 05-JAN-2001  
 by20d03.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo  
 sapiens cDNA clone by20d03 5', mRNA sequence.  
 BF727355  
 BF727355  
 BF727355.1 GI:12043266  
 EST.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 438)  
 Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.  
 NEIBANK: EST analysis and bioinformatics for ocular genomics  
 Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
 Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: Graeme@helix.nih.gov  
 Plate: 20 row: d column: 03  
 Seq primer: M13Rpl reverse primer (ABI).  
 Location/Qualifiers  
 1. 438  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="by20d03"  
 /tissue\_type="Lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens cDNA (Un-normalized, unamplified):  
 BY"  
 /notes="Organ: Eye; Vector: pCMVSPORT6; Two human lenses  
 from different adults (both approximately 40 years old)  
 together yielded 20ug of total RNA and 150ng mRNA for cDNA  
 library synthesis. A directionally cloned cDNA library in  
 the pCMVSPORT6 vector was constructed at life  
 Technologies, essentially following the protocols of the  
 SuperScript plasmid System full details of which are  
 contained in the manufacturer's instruction manual  
 (http://www.lifetech.com/). First strand synthesis was  
 carried out using a Not I primer-adaptor  
 [5'-pGACGAGTCTAGATCGGAGCGGCCG(T)15-3']. Not I/blunt  
 end inserts were cloned into the Not I/EcoR V sites in the  
 vector. EST analysis was performed on the unamplified  
 library at the NIH Intramural Sequencing Center (NISC)."

Alignment Scores:  
 Pred. No.: 3.16e-68 Length: 438  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x BF727355 (1-438)  
 QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 7 TCCTCTTCGACCGCTGCTGGACTCGGACATCTCTGAGTTCGATCCGACCGACAG 66  
 QY 21 PheValIlePheLeuAspVallyshisPheSerProGluAspLeuThrVallyshisValGln 40  
 Db 67 TTCGTCATCTTCGATGTGAAGCACTTCTCCCGAGGACCTCACCGTGAAGTGCGAG 126  
 QY 41 AspAspPheValGluIleHisGlyIlyshIsenGluArgGlnAspAspHisGlyIlyrile 60  
 Db 127 GACGACTTGTGGAGATCCACGGAAAGCACACGAGCGCCGACGACGACGACGACGACG 186

ALIGNMENTS



QY 61 SerArgGluPheHisArgGlyThrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80.  
 Db 187 TCCCGGAGTTCCACCGCGCTACCGCGCTCCACGTCGACGAGTCGGCCCTCTCT 246

QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 Db 247 TGCTCCCTGCTGCGCATGGCATGCTGACCTCTGTGGCCCAAGATCCAGACTGGCCTG 306

QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 307 GATGCCACCCAGCGGAGCGAGCCATCCCGTGTGCGGGAGGAGAGCCACCTCGGCT 366

QY 121 ProSerSer 123  
 Db 367 CCTCGTCC 375

RESULT 2  
 BM697673  
 LOCUS  
 DEFINITION  
 UI-E-DX0-agn-d-18-0-UI.r1 UI-E-DX0 Homo sapiens cDNA clone  
 UI-E-DX0-agn-d-18-0-UI 5', mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT

BM697673 446 bp mRNA linear EST 28-FEB-2002  
 UI-E-DX0-agn-d-18-0-UI.r1 UI-E-DX0 Homo sapiens cDNA clone  
 UI-E-DX0-agn-d-18-0-UI 5', mRNA sequence.

BM697673  
 EST.  
 GI:19010931  
 Homo sapiens (human)

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 446)  
 Ronaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 8889548  
 Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.

Location/Qualifiers  
 1. .446  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-DX0-agn-d-18-0-UI"  
 /tissue\_type="fetal eyes"  
 /dev\_stage="fetal"  
 /lab\_host="pH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-DX0"  
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-DX0 is a cDNA library containing the following  
 tissue(s): fetal eyes. The library was constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is AGAATCAAGA. This library  
 was created for the program, Gene Discovery in the Visual

FEATURES  
 source

ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.23e-68 Length: 446  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) \* BM697673 (1-446)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 56 TCCCTCTTCCGACCGCTGCTGACTCCGACATCTCTGAGGTTTCGATCCGACCGGCAAG 115

QY 21 PheValIlePheLeuAspVallyHisPheSerProGluAspLeuThrVallyValGln 40  
 Db 116 TTTCGTCATCTTCTCGATGTGAAGCATTCTCCCGAGGAGCTCACCGTGAAGGTGCAG 175

QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrlle 60  
 Db 176 GACGACTTTGTGGAGATCCACGGAAGACACACGAGCGCCAGACGACCCGCTACATT 235

QY 61 SerArgGluPheHisArgArgTytrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 236 TCCCGTGAGTTCCACCGCGCTACCGCTCGCTCCGCTCAACGTCGAGCCAGTCGCGCTCTCT 295

QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 Db 296 TGCTCCCTGCTGCGCATGGCATGCTGACCTTCTGTGGCCCAAGATCCAGATCCAGCTG 355

QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 356 GATGCCACCCAGCGGAGCGGAGCCATCCCGTGTGCGGGAGGAGAGGCCACCTCGGCT 415

QY 121 ProSerSer 123  
 Db 416 CCTCGTCC 424

RESULT 3  
 BF727464  
 LOCUS  
 DEFINITION  
 by21h03.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo  
 sapiens cDNA clone by21h03 5', mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)

REFERENCE  
 1 (bases 1 to 460)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 460)  
 Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.  
 NEIBANK: EST analysis and bioinformatics for ocular genomics  
 Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
 Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 21 row: h column: 03  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1. 460  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="by21h03"  
 /tissue\_type="Lens"

System, supported by National Eye Institute (NEI)."

/dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens cDNA (Un-normalized, unamplified):  
 By"  
 /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses  
 from different adults (both approximately 40 years old)  
 together yielded 20ug of total RNA and 150ng mRNA for cDNA  
 library synthesis. A directionally cloned cDNA library in  
 the pCMVSPORT6 vector was constructed at Life  
 Technologies, essentially following the protocols of the  
 Superscript Plasmid System full details of which are  
 contained in the manufacturer's instruction manual  
 (http://www.lifetech.com/). First strand synthesis was  
 carried out using a Not I primer-adaptor  
 [5'-pGACTAGTTCTAGATCGGAGCGGCCCT(T)15-3']. Not I/blunt  
 end inserts were cloned into the Not I/ScoR V sites in the  
 vector. EST analysis was performed on the unamplified  
 library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,37e-68 Length: 460  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF727464 (1-460)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 21 TCCCTCTTCGCGACCGTGGATCTCGGCATCTCTGAGTTCGATCCGCGGACAG 80  
 QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
 Db 81 TTGCTCATCTCTCGATGTAAGCACCTTCTCCCGAGGAGCCTCACCGTGAAGGTGCAG 140  
 QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60  
 Db 141 GAGCACTTTGTGGAGATCCACGAAAGCACACAGCGCGGACGAGCACCGGCTACATT 200  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 201 TCCCGTGAAGTTCACCGCGCGCTACCGCGCTCCGCGTCCACGTGGACGAGTCCGCTCT 260  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 Db 261 TGCTCCCTGTCTGCGGATGCGATGCTGACCTTCTGTGGCCCCAAGATCGAGCTGGCCTG 320  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 321 GATGCCACCCAGCGGAGCGGACCATCCCGTGTGCGGAGGAGAGAGCCACCTCGGCT 380  
 QY 121 ProSerSer 123  
 Db 381 CCTCGTCC 389

## RESULT 4

## BF727220

## LOCUS

BF727220  
 by18d12.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo  
 sapiens cDNA clone by18d12 5', mRNA sequence.

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Homo sapiens (human)  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

## AUTHORS

## TITLE

Wistow G.J., Bernstein S., Behal A. and Smith D.  
 NEIBANK: EST analysis and bioinformatics for ocular genomics

## JOURNAL

## COMMENT

Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
 Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 18 row: d column: 12  
 Seq primer: M13RF1 reverse primer (ABI).  
 Location/Qualifiers  
 1..466

## FEATURES

## source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="by18d12"  
 /tissue\_type="Lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens cDNA (Un-normalized, unamplified):  
 By"

/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses  
 from different adults (both approximately 40 years old)  
 together yielded 20ug of total RNA and 150ng mRNA for cDNA  
 library synthesis. A directionally cloned cDNA library in  
 the pCMVSPORT6 vector was constructed at Life  
 Technologies, essentially following the protocols of the  
 Superscript Plasmid System full details of which are  
 contained in the manufacturer's instruction manual  
 (http://www.lifetech.com/). First strand synthesis was  
 carried out using a Not I primer-adaptor  
 [5'-pGACTAGTTCTAGATCGGAGCGGCCCT(T)15-3']. Not I/blunt  
 end inserts were cloned into the Not I/ScoR V sites in the  
 vector. EST analysis was performed on the unamplified  
 library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,42e-68 Length: 466  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF727220 (1-466)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 29 TCCCTCTTCGCGACCGTGGATCTCGGCATCTCTGAGTTCGATCCGCGGACAG 88  
 QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
 Db 89 TTGCTCATCTCTCGATGTAAGCACCTTCTCCCGAGGAGCCTCACCGTGAAGGTGCAG 148  
 QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60  
 Db 149 GAGCACTTTGTGGAGATCCACGAAAGCACACAGCGCGGACGAGCACCGGCTACATT 208  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 209 TCCCGTGAAGTTCACCGCGCGCTACCGCGCTCCGCGTCCACGTGGACGAGTCCGCTCT 268  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 Db 269 TGCTCCCTGTCTGCGGATGCGATGCTGACCTTCTGTGGCCCCAAGATCGAGCTGGCCTG 328  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 329 GATGCCACCCAGCGGAGCGGACCATCCCGTGTGCGGAGGAGAGAGCCACCTCGGCT 388  
 QY 121 ProSerSer 123

```

Db      389 CCCTCGTCC 397

RESULT 5
BF727239
LOCUS
DEFINITION
  BF727239 500 bp mRNA linear EST 05-JAN-2001
  by18g02.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
  sapiens cDNA clone by18g02 5', mRNA sequence.
ACCESSION
BF727239
VERSION
BF727239.1 GI:12043150
KEYWORDS
EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 500)
REFERENCE
  Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
  NEIBANK: EST analysis and bioinformatics for ocular genomics
  Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
  Contact: Wistow G
  Section on Molecular Structure and Function
  National Eye Institute
  6/331, NIH, Bethesda, MD 20892-2740, USA
  Tel: 301 402 3452
  Fax: 301 496 0078
  Email: graeme@helix.nih.gov
  Plate: 18 row: g column: 02
  Seq primer: M13RPI reverse primer (ABI).
  Location/Qualifiers
    1..500
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="by18g02"
      /tissue type="Lens"
      /dev stage="Adult"
      /lab_host="EMDH108"
      /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
      BY"
      /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
      from different adults (both approximately 40 years old)
      together yielded 20ug of total RNA and 150ng mRNA for cDNA
      library synthesis. A directionally cloned cDNA library in
      the pCMVSPORT6 vector was constructed at Life
      Technologies, essentially following the protocols of the
      SuperScript Plasmid System full details of which are
      contained in the manufacturer's instruction manual
      (http://www.lifetech.com/). First strand synthesis was
      carried out using a Not I primer-adaptor
      [5'-pGACTAGTTTGTAGTCGCGAGCGCCGCC(T)15-3']. Not I/blunt
      end inserts were cloned into the Not I/EcoR V sites in the
      vector. EST analysis was performed on the unamplified
      library at the NIH Intramural Sequencing Center (NISC)."
```

ORIGIN

Alignment Scores:

Pred. No.:	3.75e-68	Length:	500
Score:	639.00	Matches:	123
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
Db:	2	Gaps:	0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF727239 (1-500)

```

Oy      1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspIys 20
Db      7 TCCTCTTCCTCCGACCGTCTGATCCCGCATCTCTGAGGTTCGATCCGACCGGACAG 66
Oy      21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIysValGln 40
Db      67 TTGTCATCTCTCTCGATGTCAGACATCTTCCCGAGGAGGACCTCACCCTGAGGTTCAG 126
Oy      41 AspAspPheValGluIleHisGlyIysHisAsnGluArgGlnAspAspHisGlyTyrIle 60

```

```

Db      127 GACGACTTTGTGGAGATCCAGCAAGCACAAACGAGCGCCAGCAGCACCGGTACATT 186
Oy      61 SerArgGluPheHisArgArgTyArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db      187 TCCCGTGAGTTCACCGCCGCTACCGCTCGCTCCCAAGTGGACAGTCGGCCCTCTCT 246
Oy      81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db      247 TGCTCCCTGTCTGCCGATGGCATGTGACCTTCTGTGGCCCCCAGATCCAGATCGGCTG 306
Oy      101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db      307 GATGCCACCCAGCGGAGGAGCCATCCCCGTGTCCGGGAGGAGAGAGACCCACCTCGGCT 366
Oy      121 ProSerSer 123
Db      367 CCCTCGTCC 375

RESULT 6
BM696489
LOCUS
DEFINITION
  BM696489 508 bp mRNA linear EST 28-FEB-2002
  UI-E-DW0-agj-n-20-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
  UI-E-DW0-agj-n-20-0-UI 5', mRNA sequence.
ACCESSION
BM696489
VERSION
BM696489.1 GI:19009747
KEYWORDS
EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 508)
REFERENCE
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
97044477
PUBMED
8889548
COMMENT
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@uiowa.edu
  Tissue Procurement: Dr. Gregg Hageman
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Researchers may obtain clones from Research
  Genetics (www.resgen.com).
  Seq primer: M13 Reverse.
  Location/Qualifiers
    1..508
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-E-DW0-agj-n-20-0-UI"
      /tissue type="lens"
      /dev stage="adult"
      /lab_host="DHI08 (Life Technologies) (T1 phage resistant)"
      /clone_lib="UI-E-DW0"
      /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
      modified polylinker; Site 1: EcoR I; Site 2: Not I;
      UI-E-DW0 is a cDNA library containing the following
      tissue(s): lens. The library was constructed according to
      Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
      1996. First strand cDNA synthesis was primed with an
      oligo-dT primer containing a Not I site. Double stranded
      cDNA was ligated to an EcoR I adaptor, digested with Not
      I, and cloned directionally into pT7T3-Pac vector. The
      oligonucleotide used to prime the synthesis of
      first-strand cDNA contains a library tag sequence that is
```

located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

## Alignment Scores:

Pred. No.: 3,836-68 Length: 508  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM696489 (1-508)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
Db 133 TCCCTCTCCGACCGTGTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGACAAG 192  
QY 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40  
Db 193 TTGCGTCATCTTCTCGATGTGAAGCAGCTTCTCCCGGAGGACCTCAACCGTGAAGGTGCAG 252  
QY 41 AspAspPheValGluIleHisGlyIleHisGlnArgGlnAspAspHisGlyTyrlle 60  
Db 253 GACGACTTTGTGGAGATCCACGGAAGCACAACGAGCGCCAGGACGACCGCTACATT 312  
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerHisValAspGlnSerAlaLeuSer 80  
Db 313 TCCCGTGAGTTCCACCGCGCTACCGCTGCGCTCCACGTGACGACGAGTCCGCGCTCTCT 372  
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
Db 373 TGCTCCCTGTGTGCGATGGCATGCTGACCTTCTGTGGCCCAAGATCCAGACTGGCTG 432  
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
Db 433 GATGCCACCCAGCGCGAGCGAGCATCCCGTGTGCGGAGGAGAGCCACCTCGGCT 492  
QY 121 ProSerSer 123  
Db 493 CCCTCGTCC 501

## RESULT 7

## BM697090

## LOCUS

DEFINITION UI-E-DW0-0-UI.r1 UI-E-DW0 Homo sapiens linear EST 28-FEB-2002  
UI-E-DW0-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone

## ACCESSION

## BM697090

## VERSION

## EST.

## SOURCE

## Homo sapiens

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

BM697090 511 bp mRNA linear EST 28-FEB-2002  
UI-E-DW0-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone  
UI-E-DW0-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone  
BM697090.1 GI:19010348  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 511)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEERF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.regen.com).  
Seq primer: M13 Reverse.

## FEATURES

## source

1..511  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-DW0-agm-k-13-0-UI"  
/tissue\_type="lens"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-DW0"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-E-DW0 is a cDNA library containing the following  
tissue(s): lens. The library was constructed according to  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was ligated to an EcoR I adaptor, digested with Not  
I, and cloned directionally into pT7T3-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is CGATTAGCGA. This library  
was created for the program, Gene Discovery in the Visual  
System, supported by National Eye Institute (NEI)."

## ORIGIN

## Alignment Scores:

Pred. No.: 3,866-68 Length: 511  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM697090 (1-511)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
Db 18 TCCCTCTCCGACCGTGTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGACAAG 77  
QY 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40  
Db .78 TTGCGTCATCTTCTCGATGTGAAGCAGCTTCTCCCGGAGGACCTCACCGTGAAGGTGCAG 137  
QY 41 AspAspPheValGluIleHisGlyIleHisGlnArgGlnAspAspHisGlyTyrlle 60  
Db 138 GACGACTTTGTGGAGATCCACGGAAGCACAACGAGCGCCAGGACGACCGCTACATT 197  
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerHisValAspGlnSerAlaLeuSer 80  
Db 198 TCCCGTGAGTTCCACCGCGCTACCGCTGCGTCCACGCTGGACCACTCGGCCCTCTCT 257  
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
Db 258 TGCTCCCTGTGTGCGATGGCATGCTGACCTTCTGTGGCCCAAGATCCAGACTGGCCTG 317  
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
Db 318 GATGCCACCCAGCGCGAGCGACCATCCCGTGTGCGGAGGAGAGAGCCACCTCGGCT 377  
QY 121 ProSerSer 123  
Db 378 CCCTCGTCC 386

## RESULT 8

## BM697099

## LOCUS

DEFINITION UI-E-DW0-0-UI.r1 UI-E-DW0 Homo sapiens linear EST 28-FEB-2002  
UI-E-DW0-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone

```

US-10-657-740-1_COPY_51_173 (1-123) x BM697099 (1-512)

ACCESSION   BM697099
VERSION     BM697099.1  GI:19010357
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 512)
AUTHORS     Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     9704477
PUBMED      889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.regen.com).
            Seq primer: M13 Reverse.

FEATURES             Location/Qualifiers
     source           1..512
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-E-DW0-aggm-m-09-0-UI"
                     /tissue_type="lens"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-DW0"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site 1: EcoR I; Site 2: Not I;
                     UI-E-DW0 is a cDNA library containing the following
                     tissue(s): lens. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CGATTAGGCA. This library
                     was created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI)."

ORIGIN
Alignment Scores:
Pred. No.:      3.87e-68      Length:      512
Score:          639.00      Matches:      123
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              4      Gaps:      0

US-10-657-740-1_COPY_51_173 (1-123) x BM697099 (1-512)

QY   1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
DB   18 TCCCTCTTCCGACCGCTGCTGACTCCGGCTCTCTGAGGTTGCTATCCGACCGGACAA 77
QY   21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40
DB   78 TTGCGTCATCTTCTCGATGTGAAGCACATTTCTCCCGGAGGACCTCACCGTGAAGGTGCAG 137

QY   41 AspAspPheValGluIleHisGlyIleHisGlyIleHisGlyIleHisGlyIleHisGlyIle 60
DB   138 GACGACTTTGTGGAGATCCACGAAAGCACACACGAGCGCCAGGACGACACGACGCTACATT 197
QY   61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
DB   198 TCCCGTGTAGTTCCACCGCGCTTACCGCTTCCGCTCAACGTCGAGACGATCGGCGCTCTCT 257
QY   81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
DB   258 TGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGGCCCCCAAGATCCAGACTGGCCTG 317
QY   101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
DB   318 GATGCCACCCAGCGGAGCGGATCCATCCCGTGTCCGCGGAGGAGGAGGAGGAGGAGGAGG 377
QY   121 ProSerSer 123
DB   378 CCTCGTCC 386

RESULT 9
LOCUS    BF726854
DEFINITION by13a10.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
sapiens cDNA clone by13a10 5', mRNA sequence.
ACCESSION BF726854
VERSION    BF726854.1  GI:12042765
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 518)
AUTHORS     Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
TITLE       NEIBANK: EST analysis and bioinformatics for ocular genomics
JOURNAL     Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 13 row: a column: 10
            Seq primer: M13RP1 reverse primer (ABI).
            Location/Qualifiers
            1..518
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="by13a10"
            /tissue_type="Lens"
            /dev_stage="Adult"
            /lab_host="EMDH10B"
            /clone_lib="Human Lens cDNA (Un-normalized, unamplified):
            By"
            /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
            from different adults (both approximately 40 years old)
            together yielded 20ug of total RNA and 150ng mRNA for cDNA
            library synthesis. A directionally cloned cDNA library in
            the pCMVSPORT6 vector was constructed at Life
            Technologies, essentially following the protocols of the
            SuperScript Plasmid System full details of which are
            contained in the manufacturer's Instruction manual
            (http://www.lifetech.com/). First strand synthesis was
            carried out using a Not I primer-adaptor
            [5'-PGACTAGTCTAGATCGGAGCGCGCTT15-3']. Not I/blunt
            end inserts were cloned into the Not I/EcoR V sites in the
            vector. EST analysis was performed on the unamplified
            library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

```

```

Alignment Scores:
Pred. No.: 3,93e-68 Length: 518
Score: 639.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x BF726854 (1-518)
QY 1 SerLeuPheAtrgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 140 TCCCTCTTCCGACCGTCTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGCAAG 199
QY 21 PheValIlePheLeuAspValIleHisGlyIleHisGlyIleHisGlyIleHisGlyIle 40
Db 200 TTCGTCACTTCTTCGATGTGAAGCATCTTCCCGGAGGACCTTCACCGTGAAGGTGCAG 259
QY 41 AspAspPheValGluIleHisGlyIleHisGlyIleHisGlyIleHisGlyIleHisGlyIle 60
Db 260 GACGACTTTGTGGAGATCCAGGAAGCACAACGAGCGCCAGGACGACGACGCTACATT 319
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 320 TCCCGTGAATTCACCGCGCGCTACCGCTCCGCTCCAACTGACGACGACGCTCTCT 379
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 380 TGCTTCCCTGTCTGCGATGTCATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGCGCTG 439
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 440 GATGCCACCCAGCCCGAGCGAGCGATCCCGTGTGCGGAGGAGAACCCACCTCGGCT 499
QY 121 ProSerSer 123
Db 500 CCCTCGTCC 508

RESULT 10
LOCUS AY419529 522 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens CRYAA gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY419529
VERSION AY419529.1 GI:39775486
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 522)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..522
Location/Qualifiers
/mol_type="genomic DNA"

/db_xref="taxon:9606"
<1..>522
/gene="CRYAA"
/locus_tag="HCM6921"

ORIGIN
Alignment Scores:
Pred. No.: 3,97e-68 Length: 522
Score: 639.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x AY419529 (1-522)
QY 1 SerLeuPheAtrgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 151 TCCCTCTTCCGACCGTCTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGCAAG 210
QY 21 PheValIlePheLeuAspValIleHisGlyIleHisGlyIleHisGlyIleHisGlyIle 40
Db 211 TTCGTCACTTCTTCGATGTGAAGCATCTTCCCGGAGGACCTTCACCGTGAAGGTGCAG 270
QY 41 AspAspPheValGluIleHisGlyIleHisGlyIleHisGlyIleHisGlyIleHisGlyIle 60
Db 271 GACGACTTTGTGGAGATCCAGGAAGCACAACGAGCGCCAGGACGACGACGCTACATT 330
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 331 TCCCGTGAATTCACCGCGCGCTACCGCTCCGCTCCAACTGACGACGACGCTCTCT 390
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 391 TGCTTCCCTGTCTGCGATGTCATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGCGCTG 450
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 451 GATGCCACCCAGCCCGAGCGAGCGATCCCGTGTGCGGAGGAGAACCCACCTCGGCT 510
QY 121 ProSerSer 123
Db 511 CCCTCGTCC 519

RESULT 11
LOCUS BF726890 523 bp mRNA linear EST 05-JAN-2001
DEFINITION by13f01.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
sapiens cDNA clone by13f01 5', mRNA sequence.
ACCESSION BF726890
VERSION BF726890.1 GI:12042801
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
TITLE Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
JOURNAL NEIBANK: EST analysis and bioinformatics for ocular genomics
COMMENT Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Place: 13 row: f column: 01
Seq primer: M13RP1 reverse primer (ABI).
FEATURES
source
1..523
Location/Qualifiers
/mol_type="mRNA"

```

/db\_xref="taxon:9606"  
 /clone="by13f01"  
 /tissue\_type="Lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH103"  
 /clone\_lib="Human Lens cDNA (Un-normalized, unamplified):  
 BY"  
 /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGATGTTTCGATCGGAGCGCGCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,98e-68 Length: 523  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF726890 (1-523)

Qy 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db TCCCTCTTCCGACCGTGTGACCTCCGGCATCTCTGAGTTCTGATCCGACCGGACCAAG 127  
 Qy 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40  
 Db TCCGTCATCTTCTCGATGTGAAGCACCTTCTCCCGGAGGACCTCACCGTGAAGGTGCAG 187  
 Qy 41 AspAspPheValGluIleHisGlyIleHisGluArgGlnAspHisGlyIleTyrIle 60  
 Db GACGACTTTGTGGAGATCCAGGAAGCACACAGAGCCGACGACGACGACGACGACGACGAC 247  
 Qy 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db TCCCGTGAATTCACCGCGCTACCGCTCCGCTCCAGTGGACGACGACGACGACGACGACGAC 307  
 Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProIleValGlnThrGlyLeu 100  
 Db TGCTCCCTGTCTGCGGATGGCATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGCGCTG 367  
 Qy 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluIleProThrSerAla 120  
 Db GATGCCACCCACCGGAGCGACCATCCCGTGTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 427  
 Qy 121 ProSerSer 123  
 Db CCCCCTGCTC 436

## RESULT 12

BM704974  
 LOCUS 533 bp mRNA linear EST 28-FEB-2002  
 DEFINITION UI-E-DW0-agg-c-09-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone  
 UI-E-DW0-agg-c-09-0-UI 5', mRNA sequence.  
 ACCESSION BM704974  
 VERSION BM704974.1 GI:19018232  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 533)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8899548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu

## COMMENT

Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.  
 Location/Qualifiers  
 1..533  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-DW0-agg-c-09-0-UI"  
 /tissue\_type="lens"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-DW0"  
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-DW0 is a cDNA library containing the following  
 tissue(s): lens. The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was ligated to an EcoR I adaptor, digested with Not  
 I, and cloned directionally into pT7T3-Pac vector. The  
 oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is CGATTAGCA. This library  
 was created for the program, Gene Discovery in the Visual  
 System, supported by National Eye Institute (NEI)."

## FEATURES

## Source

1..533  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-DW0-agg-c-09-0-UI"  
 /tissue\_type="lens"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-DW0"  
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-DW0 is a cDNA library containing the following  
 tissue(s): lens. The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was ligated to an EcoR I adaptor, digested with Not  
 I, and cloned directionally into pT7T3-Pac vector. The  
 oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is CGATTAGCA. This library  
 was created for the program, Gene Discovery in the Visual  
 System, supported by National Eye Institute (NEI)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 4.08e-68 Length: 533  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM704974 (1-533)

Qy 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db TCCCTCTTCCGACCGTGTGACCTCCGGCATCTCTGAGTTCTGATCCGACCGGACCAAG 109  
 Qy 21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40  
 Db TCCGTCATCTTCTCGATGTGAAGCACCTTCTCCCGGAGGACCTCACCGTGAAGGTGCAG 169  
 Qy 41 AspAspPheValGluIleHisGlyIleHisGluArgGlnAspHisGlyIleTyrIle 60  
 Db GACGACTTTGTGGAGATCCAGGAAGCACACAGAGCCGACGACGACGACGACGACGACGAC 229  
 Qy 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db TCCCGTGAATTCACCGCGCTACCGCTCCAGTGGACGACGACGACGACGACGACGACGAC 289



```

Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProIlyleInThrGlyLeu 100
Db 290 TGCTCCCTGTCGCCGATGCGTGTGACCTTCTGTGGCCCAAGATCCAGACTGGCCTG 349
Qy 101 AspaLThrHialaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 350 GATGCCACCCAGCCGAGGAGCCATCCCGGTGTGCGGGAGGAGAGCCACCTCGGCT 409
Qy 121 ProSerSer 123
Db 410 CCCTCGTCC 418

RESULT 13
BM705774
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1..534
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DW0-agi-h-12-0-UI"
/tissue_type="lens"
/dev_stage="adult"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DW0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-DW0 is a cDNA library containing the following
tissue(s): lens. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dr primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGATTAGCGA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."

FEATURES
source
ORIGIN

```

```

Alignment Scores:
Pred. No.: 4,09e-68 Length: 534
Score: 639.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x BM705774 (1-534)

Qy 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 78 TCCCTCTTCCGACCGCTGTGGACTCCGCGCATCTCTGAGGTTCCATCCGACCGGACAG 137
Qy 21 PheValIlePheLeuAspValIlyHisPheSerProGluAspLeuThrValLysValGln 40
Db 138 TTCGTGTCATCTTCTCGATGTGAAGCATTCTCCCGGAGGACCTCACCGTGAAGGTGCG 197
Qy 41 AspAspPheValGluIleHisGlyIlyHisAsnGluArgGlnAspAspHisGlyTyrIle 60
Db 198 GACGACTTTGTGGAGATCCACGAAAGCACACGAGCGCCAGGACGACCGGTACATT 257
Qy 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 258 TCCCGTGAAGTTCCACCGCCGCTACCGCTGCGCTCCAACTGGACCACTGGCCCTCTCT 317
Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProIlyleInThrGlyLeu 100
Db 318 TGCTCCCTGTCGCCGATGCGTGTGACCTTCTGTGGCCCAAGATCCAGACTGGCCTG 377
Qy 101 AspaLThrHialaGluArgAlaIleProValSerArgGluLysProThrSerAla 120
Db 378 GATGCCACCCAGCCGAGGAGCCATCCCGGTGTGCGGGAGGAGAGCCACCTCGGCT 437
Qy 121 ProSerSer 123
Db 438 CCCTCGTCC 446

RESULT 14
BM726399
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Wistow, G.J., Bernstein, S., Behal, A. and Smith, D.
NEIBANK: EST analysis and bioinformatics for ocular genomics
Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 06 row: a column: 09
Seq primer: M13RPI reverse primer (ABI).
Location/Qualifiers
1..536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="by06a09"
/tissue_type="lens"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Lens cDNA (Un-normalized, unamplified):"

```

BY"

/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTCTAGATCGGCGGCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,11e-68 Length: 536  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF726399 (1-536)

Qy 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 162 TCCCTCTTCCGACCGCTGCTGACTCCGGCATCTCTGAGTTTCGATCCGCGGACAA 221  
 Qy 21 PheValIlePheLeuAspVallyHisPheSerProGluAspLeuThrVallyValGln 40  
 Db 222 TTCGTTCATCTTCTCGATGTGAAGCACTTCTCCCGGAGGACCTCACCGTGAAGGTGCAG 281  
 Qy 41 AspAspPheValGluIleHisGlyIleHisGluArgGlnAspAspHisGlyTyrIle 60  
 Db 282 GACGACTTTGTGGAGATCCAGGAAGCACACAGCGCCAGGACGACCGCGGTACACTT 341  
 Qy 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 342 TCCCGTGAATTCACCGCGCTACCGCTCGCTCCAGTGGACGATCGGCGCTCTCT 401  
 Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLyLeileGlnThrGlyLeu 100  
 Db 402 TGCTCCCTGTCTCGCGATGGATGCTGACCTTCTGTGCCCCAAGATCCAGACTGGCCTG 461  
 Qy 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 462 GATGCCACCCACGCGGAGGACCATCCCGGTCTCGCGGAGGAGGAGGAGGAGGAGGAG 521  
 Qy 121 ProSerSer 123  
 Db 522 CCCTCGTCC 530

## RESULT 15

BF726679  
 LOCUS BF726679 558 bp mRNA linear EST 05-JAN-2001  
 DEFINITION by10d02.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo sapiens cDNA clone by10d02 5', mRNA sequence.  
 ACCESSION BF726679  
 VERSION BF726679.1 GI:12042590  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 558)  
 Wistow G.J., Bernstein S., Behal A. and Smith, D.  
 AUTHORS NEIBANK: EST analysis and bioinformatics for ocular genomics  
 TITLE Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
 JOURNAL Contact: Wistow G  
 COMMENT Section on Molecular Structure and Function

National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov

Plate: 10 row: d column: 02  
 Seq primer: M13RP1 reverse primer (ABI).

## FEATURES

Location/Qualifiers  
 source 1..558  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="by10d02"  
 /tissue\_type="Lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens cDNA (Un-normalized, unamplified): BY"

/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTCTAGATCGGCGGCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,33e-68 Length: 558  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF726679 (1-558)

Qy 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 140 TCCCTCTTCCGACCGCTGCTGACTCCGGCATCTCTGAGTTTCGATCCGCGGACAA 199  
 Qy 21 PheValIlePheLeuAspVallyHisPheSerProGluAspLeuThrVallyValGln 40  
 Db 200 TTCGTTCATCTTCTCGATGTGAAGCACTTCTCCCGGAGGACCTCACCGTGAAGGTGCAG 259  
 Qy 41 AspAspPheValGluIleHisGlyIleHisGluArgGlnAspAspHisGlyTyrIle 60  
 Db 260 GACGACTTTGTGGAGATCCAGGAAGCACACAGCGCCAGGACGACCGCGGTACACTT 319  
 Qy 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 320 TCCCGTGAATTCACCGCGCTACCGCTCGCTCCAGTGGACGATCGGCGCTCTCT 379  
 Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLyLeileGlnThrGlyLeu 100  
 Db 380 TGCTCCCTGTCTCGCGATGGATGCTGACCTTCTGTGCCCCAAGATCCAGACTGGCCTG 439  
 Qy 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 440 GATGCCACCCACGCGGAGGACCATCCCGGTCTCGCGGAGGAGGAGGAGGAGGAGGAG 499  
 Qy 121 ProSerSer 123  
 Db 500 CCCTCGTCC 508

## RESULT 16

BM697368  
 LOCUS UI-E-DW0-agm-d-06-0-UI.r1 569 bp mRNA linear EST 28-FEB-2002  
 DEFINITION UI-E-DW0-agm-d-06-0-UI 5', mRNA sequence.  
 ACCESSION BM697368  
 VERSION EST.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 569)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.  
 Location/Qualifiers  
 1..569  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-DW0-agm-d-06-0-UI"  
 /tissue\_type="lens"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-DW0"  
 /notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-DW0 is a cDNA library containing the following  
 tissue(s): lens. The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was ligated to an EcoR I adaptor, digested with Not  
 I, and cloned directionally into pT73-Pac vector. The  
 oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dr)18 tail. The  
 sequence tag for this library is CGATTAGCGA. This library  
 was created for the program, Gene Discovery in the Visual  
 System, supported by National Eye Institute (NEI)."

## FEATURES

source  
 1..569  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-DW0-agm-d-06-0-UI"  
 /tissue\_type="lens"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-DW0"  
 /notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-DW0 is a cDNA library containing the following  
 tissue(s): lens. The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was ligated to an EcoR I adaptor, digested with Not  
 I, and cloned directionally into pT73-Pac vector. The  
 oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dr)18 tail. The  
 sequence tag for this library is CGATTAGCGA. This library  
 was created for the program, Gene Discovery in the Visual  
 System, supported by National Eye Institute (NEI)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,448-68 Length: 569  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM697368 (1-569)

QY 21 PheValIlePheLeuAspValIysHisPheSerProGluAspLeuThrValIysValGln 40  
 Db 203 TTCTGTCATCTTCTCGATGTGAAGCACTTCTCCCGGAGGAGCTCACCGTGAAGTGCAG 262  
 QY 41 AspAspPheValGluIleHisGlyIysHisIshenGluArgGlnAspHisGlyTyrIle 60  
 Db 263 GACGACTTTGTGGAGATCCAGGAAGCACACACGAGCGCCAGCAGCCACGGCTACATT 322  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 323 TCCCGTGAGTTCACCGCGCTACCGCTCGCTCAACGAGGACCATGTGCGCCCTCTCT 382  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 Db 383 TGCTCCCTGTCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 442  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 443 GATGCCACCCACGCGAGCGAGCCATCCCGTGTGCGGAGGAGAGAGCCACCTCGCT 502  
 QY 121 ProSerSer 123  
 Db 503 CCTCGTCC 511  
 RESULT 17  
 BF726236  
 LOCUS by03f01.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo  
 DEFINITION sapiens cDNA clone by03f01 5', mRNA sequence.  
 ACCESSION BF726236  
 VERSION EST.  
 KEYWORDS BF726236.1 GI:12042147  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 576)  
 AUTHORS Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.  
 TITLE NEIBANK: EST analysis and bioinformatics for ocular genomics  
 JOURNAL Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
 COMMENT Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 03 row: f column: 01  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1..576  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="by03f01"  
 /tissue\_type="lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens cDNA (Un-normalized, unamplified):  
 BY"

## FEATURES

source  
 1..576  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="by03f01"  
 /tissue\_type="lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens cDNA (Un-normalized, unamplified):  
 BY"  
 /notes="Organ: Eye; Vector: pCMVSPORT6; Two human lenses  
 from different adults (both approximately 40 years old)  
 together yielded 20ug of total RNA and 150ng mRNA for cDNA  
 library synthesis. A directionally cloned cDNA library in  
 the pCMVSPORT6 vector was constructed at Life  
 Technologies, essentially following the protocols of the  
 SuperScript Plasmid System full details of which are  
 contained in the manufacturer's instruction manual  
 (http://www.lifetech.com/). First strand synthesis was  
 carried out using a Not I primer-adaptor  
 [5'-pGACTAGTTCGATCGAGCGGCCCT(T)15-3']. Not I/blunt  
 end inserts were cloned into the Not I/EcoR V sites in the  
 vector. EST analysis was performed on the unamplified

QY 1 SerIeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspIys 20  
 Db 143 TCCCTCTTCCGACCGCTGCTGACTCCGGCATCTCTGAGGTTCGATCCGACCGGACAG 202

library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

## Alignment Scores:

Pred. No.: 4,51e-68 Length: 576  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF726236 (1-576)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
DB 198 TCCTCTTCGCGACCGTGTGGACTCCGGCATCTCTGAGGTTGATCCGACCGGCAAG 257  
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
DB 258 TTGCTCATCTTCTCGATGTGAGCACTTCTCCCGGAGACCTCACCGTGAAGGTGCAG 317  
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60  
DB 318 GACGACTTTGTGGAGATCCACGGAAGACACACGAGCGCCAGCAGCACCGGCTACATT 377  
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
DB 378 TCCCGTGAGTTCCACCGCGCTACCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCTCT 437  
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
DB 438 TGCTCCCTGTGTGCGCATGCGATGCTGACTTCTGTGGCCCGCCAGATCCAGACTGCGCTG 497  
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
DB 498 GATGCCACCGCGCGAGGAGCCATCCCGTGTGCGGGGAGGAGAGGCCACCTCGCT 557  
QY 121 ProSerSer 123  
DB 558 CCTCGTCC 566

## RESULT 18

## BF727002

## LOCUS

DEFINITION by15c11.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo sapiens cDNA clone by15c11 5', mRNA sequence.

## ACCESSION

## BF727002

## VERSION

## EST.

## SOURCE

## Homo sapiens

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## 1 (bases 1 to 577)

## Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.

## NEIBANK: EST analysis and bioinformatics for ocular genomics

## Invest. Ophthalmol. Vis. Sci. 41 (2000) In press

## Contact: Wistow G

## Section on Molecular Structure and Function

## National Eye Institute

## 6/331, NIH, Bethesda, MD 20892-2740, USA

## Tel: 301 402 3452

## Fax: 301 496 0078

## Email: graeme@helix.nih.gov

## Plate: 15 row: c column: 11

## Seq primer: M13RPI reverse primer (ABI).

## Location/Qualifiers

## 1..577

## /organism="Homo sapiens"

## /mol\_type="mRNA"

## /db\_xref="taxon:9606"

## /clone="by15c11"

## /tissue\_type="Lens"

## FEATURES

## source

## RESULT 19

## BF721893

## LOCUS

## DEFINITION

## UI-E-E00-ahw-c-15-0-UI.r1

## UI-E-E00 Homo sapiens

## cDNA clone

## 5', mRNA sequence.

## ACCESSION

## BF721893

## VERSION

## EST.

## KEYWORDS

## SOURCE

## Homo sapiens

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## 1 (bases 1 to 578)

## Bonaldo,M.F., Lennon,G. and Soares,M.B.

## Normalization and subtraction: two approaches to facilitate gene

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.reagen.com).  
Seq primer: M13 Reverse.

FEATURES  
source

Location/Qualifiers  
1..578  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-E00-ahw-c-15-0-UI"  
/tissue\_type="fetal eye"  
/dev\_stage="fetal"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-E00"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-E00 is a cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Alignment Scores:  
Pred. No.: 4,53e-68 Length: 578  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM721893 (1-578)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
Db 205 TCCTCTTCGACACCGTGTGGACTCCGGCATCTCTGAGGTTGATCCGACCGGACAAG 264  
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
Db 265 TTCGTATCTTCCTCGATGTGAAGCACTTCTCCCGGAGGACCTCACCGTGAAGTGCAG 324  
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60  
Db 325 GAGCACTTTGTGAGATCCACGAAGACACACGACGCGCAGCAGCAGCGGTACATT 384  
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
Db 385 TCCCGTGAGTTCCACCGCGCTACCGCTGCGGTCCACGTGGACCATCGCCCTCTCT 444  
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100

Db 445 TGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCCCAAGATCCAGACTGCGCTG 504  
QY 101 AspAlaThrHisAlaGluAlaGluAlaProValSerArgGluGluLysProThrSerAla 120  
Db 505 GATGCCACCCACCCCGGAGCGCCATCCCGTGTGCGGGAGGAGAGCCACCTCGGCT 564  
QY 121 ProSerSer 123  
Db 565 CCCTCGTCC 573

RESULT 20  
LOCUS

BF727295

By19e10.y1 Human Lens cDNA (Un-normalized, mRNA linear EST 05-JAN-2001 sapiens cDNA clone by19e10 5', mRNA sequence).

ACCESSION

BF727295

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Wistow.G.J., Bernstein.S., Behal.A. and Smith.D.

NEIBANK: EST analysis and bioinformatics for ocular genomics

Invest. Ophthalmol. Vis. Sci. 41 (2000) In press

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 19 row: e column: 10

Seq primer: M13RP1 reverse primer (ABI).

Location/Qualifiers

1..592

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="by19e10"

/tissue\_type="Lens"

/dev\_stage="Adult"

/lab\_host="EMDH10B"

/clone\_lib="Human Lens cDNA (Un-normalized, unamplified):

By"

/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses

from different adults (both approximately 40 years old)

together yielded 20ug of total RNA and 150ng mRNA for cDNA

library synthesis. A directionally cloned cDNA library in

the pCMVSPORT6 vector was constructed at life

Technologies, essentially following the protocols of the

SuperScript Plasmid System full details of which are

contained in the manufacturer's instruction manual

(http://www.lifetech.com/). First strand synthesis was

carried out using a Not I primer-adaptor

[5'-pGACTAGTCTAGATCGGCGCGCC(T)15-3']. Not I/blunt

end inserts were cloned into the Not I/EcoR V sites in the

vector. EST analysis was performed on the unamplified

library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:  
Pred. No.: 4,68e-68 Length: 592  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF727295 (1-592)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20

```

Db      161 TCCCTCTTCGACCGCTGCTGACTCCGGCATCTCTGAGGTTGATCCCGAGGACAAG 220
QY      21  PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40
Db      221 TTCTGTCATCTCTCTCGATGTGAAGCACTTCTCCCGAGGACCTCAACCGTGAAGGTGCAG 280
QY      41  AspAspPheValIleHisGlyHisGlyHisAsnGluArgGlnAspHisGlyTyrIle 60
Db      281 GACGACTTTGTGTGAGATCCACGAAAGCAACAGAGCGCCGACGACGACCGCTACATT 340
QY      61  SerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db      341 TCCCGTGAATTCACCGCGCTACCGCTGCGCTGCTCCAGTCCAGCGCGCCCTCTCT 400
QY      81  CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProIleGlnThrGlyLeu 100
Db      401 TGTCTCTCTGTCGCGATGCGATGCTGACTTCTGTGGCCCCCAAGATCCAGACTGGCCCTG 460
QY      101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db      461 GATGCCACCCAGCGCGAGCGACCATCCCGTGTCCGGGAGGAGAGCCACCTCGGCT 520
QY      121 ProSerSer 123
Db      521 CCTCGTCC 529

BM705926      592 bp      mRNA      linear      EST 28-FEB-2002
UI-E-DW0-agh-f-18-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
UI-E-DW0-agh-f-18-0-UI 5', mRNA sequence.
ACCESSION      BM705926
VERSION        1
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          1 (bases 1 to 592)
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL        Genome Res. 6 (9), 791-806 (1996)
MEDLINE        97044477
PUBMED         8889548
COMMENT        Contact: Soares, MB
                Coordinated Laboratory for Computational Genomics
                University of Iowa
                375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
                Tel: 319 335 8250
                Fax: 319 335 9565
                Email: bento-soares@uiowa.edu
                Tissue Procurement: Dr. Gregg Hageman
                cDNA library preparation by: Dr. M. Bento Soares, University of Iowa
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: Researchers may obtain clones from Research
                Genetics (www.resgen.com).
                Seq primer: M13 Reverse.
                Location/Qualifiers
                1..592
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="UI-E-DW0-agh-f-18-0-UI"
                /tissue_type="lens"
                /dev_stage="adult"
                /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                /clone_lib="UI-E-DW0"
                /note="Organ: eye; Vector: p773-Pac (Pharmacia) with a
                modified polylinker; Site 1: EcoR I; Site 2: Not I;
                UI-E-DW0 is a cDNA library containing the following

```

tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,68e-68 Length: 592  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM705926 (1-592)

```

QY      1  SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db      202 TCCCTCTTCGACCGCTGCTGACTCCGGCATCTCTGAGGTTGATCCCGAGGACAAG 261
QY      21  PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40
Db      262 TTCTGTCATCTCTCTCGATGTGAAGCACTTCTCCCGAGGACCTCAACCGTGAAGGTGCAG 321
QY      41  AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspHisGlyTyrIle 60
Db      322 GACGACTTTGTGAGATCCACGAAAGCAACAGAGCGCCGAGCGACCGCTACATT 381
QY      61  SerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db      382 TCCCGTGAATTCACCGCGCTACCGCTGCGCTCCAGCTGGACCGACCGCTCTCT 441
QY      81  CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProIleGlnThrGlyLeu 100
Db      442 TCTCTCTCTGTGCGGATGGCATGCTTCTGTGGCCCCCAAGATCCAGACTGGCCTG 501
QY      101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db      502 GATGCCACCCAGCGCGAGCGACCATCCCGTGTCCGGGAGGAGAGCCACCTCGGCT 561
QY      121 ProSerSer 123
Db      562 CCTCGTCC 570

```

## RESULT 22

BM697101

LOCUS

DEFINITION

UI-E-DW0-agh-m-13-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone

UI-E-DW0-agh-m-13-0-UI 5', mRNA sequence.

ACCESSION

BM697101

VERSION

BM697101.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 593)

AUTHORS

Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

8889548

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa  
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
Seq primer: M13 Reverse.

## FEATURES

Location/Qualifiers  
1. .596  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-DW0-agg-m-13-0-UI"  
/dev\_stage="adult"  
/tissue\_type="lens"  
/lab\_host="DHI0B (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

## Alignment Scores:

Pred. No.: 4,696-68 Length: 593  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM697101 (1-593)

Qy	1	SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys	20
Db	136	TCCTCTTCGCGACCGTGTGACTCCGGCACTCTGAGGTTCGATCCGACCGGACAA	195
Qy	21	PheValIlePheLeuAspValLysHisSphSerProGluAspLeuThrValLysValGln	40
Db	196	TTCGTCATCTTCCTCGATGTGAAGCACTTCTCCCGAGGACCTCACCCTGAGGTGAG	255
Qy	41	AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle	60
Db	256	GACGACTTTGTGGAGATCCACGAAGACACACGAGCGCCAGGACGACACCGGTACAT	315
Qy	61	SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer	80
Db	316	TCCTGTAGTTCCACCGCGCTTACCGCTGCGTCCACGTGGACCATCGGCTCTCT	375
Qy	81	CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu	100
Db	376	TGCTTCCTGTCTGCGATGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	435
Qy	101	AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla	120
Db	436	GATGCCACCCACCGCGAGGCCATCCCTGTCTGCGGGAGGAGAGGCCACCTCGGCT	495

Qy 121 ProSerSer 123  
Db 496 CCCTCGTCC 504

## RESULT 23

## LOCUS

## DEFINITION

BM696581 596 bp mRNA linear EST 28-FEB-2002  
UI-E-DW0-agg-o-15-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone  
UI-E-DW0-agg-o-15-0-UI 5', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

## CONTACT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## Alignment Scores:

Pred. No.: 4,726-68 Length: 596  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM696581 (1-596)



QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 DB 206 TCCCTCTTCGCGACCGTGTGACATCCGCGATCTCTGAGGTTCGATCCGACCGGACCAAG 265  
 QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
 DB 266 TTCGTCACTTCCTCGATGTGAGCACTTCCTCCCGAGGAGCACTCACCGTGAAGTGTGAG 325  
 QY 41 AspAspPheValGluIleHisGlyLysHisAenGluArgGlnAspAspHisGlyTyrIle 60  
 DB 326 GACGACTTGTGAGATCCACGGAAGCACACGAGCGCCAGGACGACGACCGCTACATTT 385  
 QY 51 SerArgGluPheHisArgArgTyrArgLeuProSerAenValAspGlnSerAlaLeuSer 80  
 DB 386 TCCCGTGAAGTTCCACGCGCGCTACCGCTCCGCTCCAACTGAGGACGACGCTGCTCTCT 445  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 DB 446 TGCTCCCTGTTCGCGATGCGATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGGCCTG 505  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 DB 506 GATGCCACCCAGCGCGAGGACCATCCCGTGTGCGGGAGGAGAGCCCACTCGGCT 565  
 QY 121 ProSerSer 123  
 DB 566 CCCTCGTCC 574

RESULT 24  
 BM696976 599 bp mRNA linear EST 28-FEB-2002  
 LOCUS UI-E-DW0-agl-f-04-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone  
 DEFINITION UI-E-DW0-agl-f-04-0-UI 5', mRNA sequence.  
 ACCESSION BM696976  
 VERSION  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 599)  
 REFERENCE  
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 9704477  
 PUBMED 889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.  
 Location/Qualifiers  
 1: 599  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-DW0-agl-f-04-0-UI"  
 /tissue\_type="lens"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-DW0"  
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-DW0 is a cDNA library containing the following  
 tissue(s): lens. The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was ligated to an EcoR I adaptor, digested with Not  
 I, and cloned directionally into pT7T3-Pac vector. The  
 oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is CGATTAGCGA. This library  
 was created for the program, Gene Discovery in the Visual  
 System, supported by National Eye Institute (NEI)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,75e-68 Length: 599  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM696976 (1-599)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 DB 70 TCCCTCTTCGCGACCGTGTGACATCCGCGATCTCTGAGGTTCGATCCGACCGGACCAAG 129  
 QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
 DB 130 TTCGTCACTTCCTCGATGTGAGCACTTCCTCCCGAGGAGCACTCACCGTGAAGTGTGAG 189  
 QY 41 AspAspPheValGluIleHisGlyLysHisAenGluArgGlnAspAspHisGlyTyrIle 60  
 DB 190 GACGACTTGTGAGATCCACGGAAGCACACGAGCGCCAGGACGACGACCGCTACATTT 249  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAenValAspGlnSerAlaLeuSer 80  
 DB 250 TCCCGTGAAGTTCCACCGCGCTACCGCTGCGCTCCAACTGAGGACGACGCTGCTCTCTCT 309  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 DB 310 TGCTCCCTGTTCGCGATGCGATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGGCCTG 369  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 DB 370 GATGCCACCCAGCGCGAGGACCATCCCGTGTGCGGGAGGAGAGCCCACTCGGCT 429  
 QY 121 ProSerSer 123  
 DB 430 CCCTCGTCC 438

## RESULT 25

CD675250

LOCUS

DEFINITION

fs21c02.v1 Human Lens cDNA (Normalized): fs Homo sapiens cDNA clone

fs21c02.5', mRNA sequence.

ACCESSION

CD675250

VERSION

CD675250.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 607)

AUTHORS

Wistow,G., Bernstein,S.I., Wyatt,M.K., Behal,A., Touchman,J.W.,

Bouffard,G., Smith,D. and Peterson,K.

TITLE

Expressed sequence tag analysis of adult human lens for the NEIBank

Project: over 2000 non-redundant transcripts, novel genes and

splice variants

JOURNAL

Mol. Vis. 8 (4), 171-184 (2002)

MEDLINE 22103463  
 PUBMED 12107413  
 COMMENT Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 21 row: c column: 02  
 Seq primer: M13RPI reverse primer (ABI).  
 Location/Qualifiers  
 1. .607  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="fs21c02"  
 /tissue\_type="lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens CDNA (Normalized): fs"  
 /note="Organ: Eye; Vector: pCMVSPORT6; A human lens library (by) was normalized by self-subtraction. One portion of double stranded plasmid DNA representing the library was linearized by NotI. This NotI digested library was used as a template for biotinylated RNA synthesis using SP6 RNA polymerase. Another portion of the double stranded plasmid library was converted to single-stranded circles in vitro using Gene II and Exonuclease III (Life Technologies). Single-stranded DNA (1 mg) was hybridized (Cot 500) with 41 mg of Bio-RNA and vector blocking oligonucleotides. The hybridized Bio-RNA/ss-circles were removed by streptavidin:phenol extraction. EST analysis was performed on the library at the NIH Intramural Sequencing Center (NISC)."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,83e-68 Length: 607  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservatative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x CD675250 (1-607)  
 QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 DB 201 TCCCTCTCCGACCGCTGCTGACTCCGGCATCTCTGAGTTTCGATCCGACCGGACAG 260  
 QY 21 PheValIlePheLeuAspVallyshisPheSerProGluAspLeuThrValLysValGln 40  
 DB 261 TTCGTATCTTCTCGATGTGAAGACATCTTCCCGGAGGACCTCACCGTGAAGGTGCAG 320  
 QY 41 AspAspPheValGluIleHisGlyIlyshisAsnGluArgGlnAspAspHisGlyTyrIle 60  
 DB 321 GACGACTTGTGAGATCCAGGAAGCACACAGCGCCAGGACGACCGGTATCAT 380  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 DB 381 TCCCGTAGTTTCCACCGCGGTACCGCTGCGCTCCACGTGGACCATGCGCCCTCTCT 440  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProTylsIleGlnThrGlyLeu 100  
 DB 441 TGCTCCCTGTCTGCGATGGATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGGCCTG 500  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 DB 501 GATGCCACCCAGCGGACGAGCCATCCCGTGTGCGGGAGGAGGAGCCACCTCGGCT 560  
 QY 121 ProSerSer 123  
 |||||

Db 561 CCCTCGTCC 569  
 RESULT 26  
 LOCUS CD676130  
 DEFINITION fs32e02.y1 Human Lens cDNA (Normalized): fs Homo sapiens cDNA clone  
 fs32e02 5', mRNA sequence.  
 ACCESSION CD676130  
 VERSION CD676130  
 KEYWORDS EST.  
 SOURCE CD676130.1 GI:32177861  
 ORGANISM Homo sapiens (human)  
 REFERENCE 1 (bases 1 to 612)  
 AUTHORS Wistow,G., Bernstein,S.L., Wyatt,M.K., Behal,A., Touchman,J.W.,  
 Bouffard,G., Smith,D. and Peterson,K.  
 TITLE Expressed sequence tag analysis of adult human lens for the NEIBank  
 Project: over 2000 non-redundant transcripts, novel genes and  
 splice variants  
 JOURNAL Mol. Vis. 8 (4), 171-184 (2002)  
 MEDLINE 22103463  
 PUBMED 12107413  
 COMMENT Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 32 row: e column: 02  
 Seq primer: M13RPI reverse primer (ABI).  
 Location/Qualifiers  
 1. .612  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="fs32e02"  
 /tissue\_type="lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens CDNA (Normalized): fs"  
 /note="Organ: Eye; Vector: pCMVSPORT6; A human lens library (by) was normalized by self-subtraction. One portion of double stranded plasmid DNA representing the library was linearized by NotI. This NotI digested library was used as a template for biotinylated RNA synthesis using SP6 RNA polymerase. Another portion of the double stranded plasmid library was converted to single-stranded circles in vitro using Gene II and Exonuclease III (Life Technologies). Single-stranded DNA (1 mg) was hybridized (Cot 500) with 41 mg of Bio-RNA and vector blocking oligonucleotides. The hybridized Bio-RNA/ss-circles were removed by streptavidin:phenol extraction. EST analysis was performed on the library at the NIH Intramural Sequencing Center (NISC)."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,88e-68 Length: 612  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservatative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-657-740-1\_COPY\_51\_173 (1-123) x CD676130 (1-612)  
 QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 DB 6 TCCCTCTCCGACCGCTGCTGACTCCGGCATCTCTGAGTTTCGATCCGACCGGACAG 65  
 QY 21 PheValIlePheLeuAspVallyshisPheSerProGluAspLeuThrValLysValGln 40

```

Db      66  TTGTCATCTTCTCGATGTGAGACATTTCTCCCGAGGAGCCTCACCGTGAAGTGCAG 125
Qy      41  AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60
Db      126  GACGACTTTGTGGAGATCCACGGAAGCACAAACAGAGCCAGGACGACCCAGCTTACATT 185
Qy      61  SerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db      186  TCCCGTAGTTCACACCGCGCTTACCGCTCCAAACGTCGACGACGTCGCGCCCTCTCT 245
Qy      81  CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db      246  TGCTCCCTGTCTGCGATGGCATGTGACCTTCTGTGCCCCCAGATCCAGACTGGCCTG 305
Qy      101  AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db      306  GATGCCACCCACGCGAGCGAGCCATCCCGTGTGCGGGAGGAGAACCCACCTCGGCT 365
Qy      121  ProSerSer 123
Db      366  CCCTCGTCC 374

```

```

RESULT 27
BF727028
LOCUS   BF727028
DEFINITION by15905.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo sapiens cDNA clone by15905 5', mRNA sequence.
ACCESSION BF727028
VERSION   BF727028.1 GI:12042939
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 629)
Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
TITLE    NEIBANK: EST analysis and bioinformatics for ocular genomics
JOURNAL  Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT  Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 15 row: g column: 05
Seq primer: M13RP1 reverse primer (ABI).

```

## FEATURES

```

source
1..629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="by15905"
/tissue_type="Lens"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Lens cDNA (Un-normalized, unamplified): BY"
/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTTAGATCGAGCGCGCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."
```

## ORIGIN

```

Alignment Scores:
Pred. No.: 5,06e-68 Length: 629
Score: 639.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x BF727028 (1-629)
Qy      1  SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db      198  TCCCTCTTCCGACCGTGTGAGCTCCGCGATCTCTGAGGTTCGATCGACCGGACAAAG 257
Qy      21  PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40
Db      258  TTGTCATCTTCTCGATGTGAGACATTTCTCCCGAGGAGCCTCACCGTGAAGTGCAG 317
Qy      41  AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60
Db      318  GACGACTTTGTGGAGATCCACGGAAGCACAAACAGAGCCAGGACGACCCAGCTTACATT 377
Qy      61  SerArgGluPheHisArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db      378  TCCCGTAGTTCACACCGCGCTTACCGCTCCAAACGTCGACGACCTCGGCGCTCTCT 437
Qy      81  CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db      438  TGCTCCCTGTCTGCGATGGCATGTGACCTTCTGTGCCCCCAGATCCAGACTGGCCTG 497
Qy      101  AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db      498  GATGCCACCCACGCGAGCGAGCCATCCCGTGTGCGGGAGGAGAACCCACCTCGGCT 557
Qy      121  ProSerSer 123
Db      558  CCCTCGTCC 566

```

## RESULT 28

```

LOCUS   BM696651
DEFINITION UI-E-DW0-agk-k-24-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
ACCESSION BM696651
VERSION   BM696651.1 GI:19009909
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 630)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE    Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL  Genome Res. 6 (9), 791-806 (1996)
MEDLINE  97044477
PUBMED  8889548
COMMENT  Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.

```

FEATURES  
source

Location/Qualifiers  
1. .631  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-DW0-agh-k-24-0-UI"  
/tissue\_type="lens"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-DW0"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

Alignment Scores:  
Pred. No.: 5,07e-68 Length: 630  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM696651 (1-630)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
Db 204 TCCCTCTTCGCGACCGCTGGACTCCGGCATCTCTGAGGTTCGATCCGCGGACCAAG 263  
QY 21 PheValIlePheLeuAspVallyHisPheSerProGluAspLeuThrVallyValGln 40  
Db 264 TTGCTCATCTTCTCGATGTGAAGCACTTCTCCCGGAGGACCTCACCGTGAAGTGCAG 323  
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspHisGlyTyrIle 60  
Db 324 GACGACTTTGTGAGATCCACGGAAAGCACACGAGCGCGAGGACGACCGGCTACATT 383  
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
Db 384 TCCGTGAGTTCCACCGCGCTACCGCTGCGCTCCCAACGTCGAGGACGACGCGCTCTCT 443  
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
Db 444 TGCTCCCTGTCTGCGGATGGCATGTGACCTTCTGTGACCCCAAGATCCAGCTGGCGTG 503  
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
Db 504 GATGCCACCCAGCGGAGGACCATCCCGTGTGCGGGAGGAGAGGCCACCTCGGCT 563  
QY 121 ProSerSer 123  
Db 564 CCCTCGTCC 572

RESULT 29  
BM706251  
LOCUS  
DEFINITION  
UI-E-DW0-agh-g-05-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone  
UI-E-DW0-agh-g-05-0-UI 5', mRNA sequence.  
ACCESSION  
BM706251  
VERSION  
BM706251.1 GI:19019509  
KEYWORDS  
EST.

SOURCE  
ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

97044477  
889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Reverse.

FEATURES  
source

Location/Qualifiers  
1. .631  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-DW0-agh-g-05-0-UI"  
/tissue\_type="lens"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-DW0"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.08e-68 Length: 631  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM706251 (1-631)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
Db 50 TCCCTCTTCGCGACCGCTGGACTCCGGCATCTCTGAGGTTCGATCCGCGGACCAAG 109  
QY 21 PheValIlePheLeuAspVallyHisPheSerProGluAspLeuThrVallyValGln 40  
Db 110 TTGCTCATCTTCTCGATGTGAAGCACTTCTCCCGGAGGACCTCACCGTGAAGTGCAG 169  
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspHisGlyTyrIle 60  
Db 170 GACGACTTTGTGAGATCCACGGAAAGCACACGAGCGCGGACGACGCTACATT 229

```

QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
    |||||
Db 230 TCCCTGTGAGTTCACCGCGCTACCGCTCGCGTCCAGCTGGACCAAGTCGCGCTCTCT 289
    |||||
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
    |||||
Db 290 TGCTCCCTGTCTGCCGATGGCATGTGACCTTCTGTGGCCCCCAAGATCCAGACTGCCCTG 349
    |||||
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
    |||||
Db 350 GATGCCACCCAGCGCGAGCGACCATCCCGTGTGCGGGAGGAGAGGCCACCTCGGCT 409
    |||||
QY 121 ProSerSer 123
    |||||
Db 410 CCTCGTCC 418

RESULT 30
CD672144
LOCUS
DEFINITION
    fg10b07.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
    fg10b07 5', mRNA sequence.
ACCESSION
    CD672144
VERSION
    CD672144.1 GI:32173875
KEYWORDS
    EST.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 633)
AUTHORS
    Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
    Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE
    Expressed sequence tag analysis of adult human iris for the NEIBank
    Project: steroid-response factors and similarities with retinal
    pigment epithelium
JOURNAL
    Mol. Vis. 8 (4), 185-195 (2002)
MEDLINE
    22103462
PUBMED
    12107412
COMMENT
    Contact: Wistow G
    Section on Molecular Structure and Function
    National Eye Institute
    6/331, NIH, Bethesda, MD 20892-2740, USA
    Tel: 301 402 3452
    Fax: 301 496 0078
    Email: graeme@helix.nih.gov
    Plate: 10 row: b column: 07
    Seq primer: M13Rpl reverse primer (ABI).
FEATURES
    source
    1..633
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="fg10b07"
        /tissue_type="iris"
        /dev_stage="Adult"
        /lab_host="EMDH10B"
        /clone_lib="Human Iris cDNA (Normalized): fg"
        /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
        library (bx) was normalized by self-subtraction. One
        portion of double stranded plasmid DNA representing the
        library was linearized by NotI. This NotI digested library
        was used as a template for biotinylated RNA synthesis
        using SP6 RNA polymerase. Another portion of the double
        stranded plasmid library was converted to single-stranded
        circles in vitro using Gene II and Exonuclease III (Life
        Technologies). Single-stranded DNA (1 mg) was hybridized
        (Cot 500) with 41 mg of Bio-RNA and vector blocking
        oligonucleotides. The hybridized Bio-RNA/ss-circles were
        removed by streptavidin:phenol extraction. EST analysis
        was performed on the library at the NIH Intramural
        Sequencing Center (NISC)."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 5,1e-68 Length: 633
Score: 639.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-657-740-1_COPY_51_173 (1-123) x CD672144 (1-633)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
    |||||
Db 202 TCCCTCTTCCGACCGCTGCTGGACTCCGGCATCTCTGAGTTCCATCCAGCCGGACAG 261
    |||||
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40
    |||||
Db 262 TTGCTCATCTTCTCGATGTGAAGCACTTCTCCCGGAGGACCTCACCGTGAAGGTGCAG 321
    |||||
QY 41 AspAspPheValGluIleHisGlyLysHisGlnGluArgGlnAspAspHisGlyTyrIle 60
    |||||
Db 322 GACGACTTTGTGGAGATCCAGGAAGCACAAACGAGCGCCAGACGACCCACCGCTACATT 381
    |||||
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
    |||||
Db 382 TCCCGTGAGTTCCACCGCGCTACCGCTCGCTCCAGCTGACGACGACGACGACGCTCTCT 441
    |||||
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
    |||||
Db 442 TGCTCCCTGTCTGCCGATGGCATGTGACCTTCTGTGGCCCCCAAGATCCAGACTGCCCTG 501
    |||||
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
    |||||
Db 502 GATGCCACCCAGCGCGAGCGACCATCCCGTGTGCGGGAGGAGAGGCCACCTCGGCT 561
    |||||
QY 121 ProSerSer 123
    |||||
Db 562 CCTCGTCC 570

RESULT 31
BM697066
LOCUS
DEFINITION
    UI-E-DW0-agm-g-03-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
    UI-E-DW0-agm-g-03-0-UI 5', mRNA sequence.
ACCESSION
    BM697066
VERSION
    BM697066.1 GI:19010324
KEYWORDS
    EST.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 635)
AUTHORS
    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE
    Normalization and subtraction: two approaches to facilitate gene
    discovery
JOURNAL
    Genome Res. 6 (9), 791-806 (1996)
MEDLINE
    9704477
PUBMED
    8889548
COMMENT
    Contact: Soares, MB
    Coordinated Laboratory for Computational Genomics
    University of Iowa
    375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
    Tel: 319 335 8250
    Fax: 319 335 9565
    Email: bento-soares@uiowa.edu
    Tissue Procurement: Dr. Gregg Hageman
    cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
    cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
    DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
    Clone Distribution: Researchers may obtain clones from Research
    Genetics (www.resgen.com).
    Seq primer: M13 Reverse.
FEATURES
    Location/Qualifiers
    1..635
    source

```

FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cloned="UI-E-DW0-agh-g-03-0-UI"
/tissue_type="lens"
/dev_stage="adult"
/lab_host="DHI0B (Life Technologies) (T1 phage resistant)"
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 5,12e-68 Length: 635
Score: 639.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM697066 (1-635)

```

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 148 TCCTCTTCGCGACCGTGTGAGCTCCGGATCTCTGGAGTTCTGATCCGACCGGACAA 207
QY 21 PheValIlePheLeuAspValIleHisGlyLysHisAsnGluArgGlnAspHisGlyTyrIle 60
Db 208 TTGCTGATCTTCCTCGATGTGAGCACTTCTCCCGAGGAGACTTACCGTGAAGTGCGAG 267
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspHisGlyTyrIle 60
Db 268 GAGCACTTTGTGAGATCCACGGAAGACAAACGACGCGGACGACGACGCGCTACATT 327
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db 328 TCCCGTGAGTTCCACCGCGCTACCGCTCCGCTCCAGCTGGACCACTGCGGCTCTCT 387
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db 388 TGCTCTCTCTCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db 448 GATGCCACCCACCGGAGGAGCCATCCCGTGTGCGGCGGAGGAGGAGGAGGAGGAGGAG 507
QY 121 ProSerSer 123
Db 508 CCCTCGTCC 516

```

```

RESULT 32
BM706270
LOCUS
DEFINITION
UI-E-DW0-agh-k-09-0-UI.r1 UI-E-DW0 Homo sapiens linear EST 28-FEB-2002
VERSION
BM706270
ACCESSION
BM706270.1 GI:19019528
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

```

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 659)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
Seq primer: M13 Reverse.

## FEATURES

## source

```

Location/Qualifiers
1..659
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cloned="UI-E-DW0-agh-k-09-0-UI"
/tissue_type="lens"
/dev_stage="adult"
/lab_host="DHI0B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DW0"
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTAGCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 5,38e-68 Length: 659
Score: 639.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM706270 (1-659)

```

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db 207 TCCTCTTCGCGACCGTGTGAGCTCCGGATCTCTGGAGTTCTGATCCGACCGGACAA 266
QY 21 PheValIlePheLeuAspValIleHisGlyLysHisAsnGluArgGlnAspHisGlyTyrIle 40
Db 267 TTGCTGATCTTCCTCGATGTGAGCACTTCTCCCGAGGAGACTTACCGTGAAGTGCGAG 326
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspHisGlyTyrIle 60
Db 327 GAGCACTTTGTGAGATCCACGGAAGACAAACGACGCGGACGACGACGCTACATT 386
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80

```

```

Db      387 TCCCTGAGTTCACCGCGCTACCGCTCGCTCCGCTCAACGTGACCAAGTCGCGCTCTCT 446
QY      81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db      447 TGCTCCCTGTCTGCGGATGCTGACCTTCTGTGGCCCCCAAGATCCAGACTGCGCTG 506
QY      101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db      507 GATGACCAACCCAGCGGAGCGGATCCCGCTGTCGCGGAGGAGGAGGCCACCTCGGCT 566
QY      121 ProSerSer 123
Db      567 CCCTCGTCC 575

RESULT 33
BM697160
LOCUS      661 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-DW0-9gm-i-06-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
VERSION     UI-E-DW0-9gm-i-06-0-UI 5', mRNA sequence.
ACCESSION   BM697160.1 GI:19010418
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 661)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     9704477
PUBMED      8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES    Location/Qualifiers
             1..661
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="UI-E-DW0-9gm-i-06-0-UI"
                /tissue_type="lens"
                /dev_stage="adult"
                /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                /clone_lib="UI-E-DW0"
                /note="organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                modified polylinker; Site_1: EcoR I; Site_2: Not I;
                UI-E-DW0 is a cDNA library containing the following
                tissue(s): lens. The library was constructed according to
                Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                1996. First strand cDNA synthesis was primed with an
                oligo-dT primer containing a Not I site. Double stranded
                cDNA was ligated to an EcoR I adaptor, digested with Not
                I, and cloned directionally into pT7T3-Pac vector. The
                oligonucleotide used to prime the synthesis of
                first-strand cDNA contains a library tag sequence that is
                located between the Not I site and the (dT)18 tail. The
                sequence tag for this library is CGATTAGGCA. This library
                was created for the program, Gene Discovery in the Visual
                System, supported by National Eye Institute (NEI)."
```

## ORIGIN

## Alignment Scores:

5.4e-68 Length: 661  
 Score: 639.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM697160 (1-661)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 Db 144 TCCCTCTTCCGCGACCGCTGCTGACCTCCGCGCATCTCTGAGGTTTCGATCCGACCGGCAAG 203  
 QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
 Db 204 TTCGTCATCTTCTCGATGTGAGCACCTTCTCCCGGAGGACCTTACCGTGAAGGTGCAG 263  
 QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60  
 Db 264 GACGACTTGTGGAGATCCACGAAAGCACACAGAGCGCCAGCAGCCACCGCTACATT 323  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 Db 324 TCCCGTGAAGTTCCACCGCGCTACCGCGCTCCGCGTCCAAAGTGGAGCAGTCGCGCTCTCT 383  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
 Db 384 TGCTCCCTGTCTGCGGATGGCATGCTGACCTTCTGTGCCCCCAGATCCAGATCGGCTG 443  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
 Db 444 GATGCCACCCAGCGGAGCGAGCCATCCCGTGTGCGGAGGAGAGCCACCTCGGCT 503  
 QY 121 ProSerSer 123  
 Db 504 CCCTCGTCC 512

## RESULT 34

## BX118596

## LOCUS

## DEFINITION

## IMAGp9898P14436 ; IMAGE:220285, mRNA sequence.

## ACCESSION

## BX118596

## VERSION

## BX118596.1 GI:27882169

## KEYWORDS

## EST.

## SOURCE

## Homo sapiens (human)

## ORGANISM

## Homo sapiens

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

## 1 (bases 1 to 681)

## AUTHORS

## Radelof,U., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,

## TITLE

## Human Unigeneset - RZPD3

## JOURNAL

## Unpublished (2003)

## COMMENT

## Contact: Ina Rolfs

## RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

## Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

## RZPD; IMAGp9898P14436.

## RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

## Human Unigeneset - RZPD3 (RZPDLIB No.972)

## http://www.rzpd.de/CloneCards/cgi-

## bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs

## RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

## Heubnerweg 6, D-14059 Berlin, Germany

## Tel: +49 30 32639 101

## Fax: +49 30 32639 111

## www.rzpd.de

## This clone is available royalty-free from RZPD;

## contact RZPD (clone@rzpd.de) for further information. Seq primer:

## M13r, Primer sequence: TTTCCACAGGAAACAGCTATGAC.



FEATURES  
source

Location/Qualifiers  
1. .681  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE998P14436 ; IMAGE:220285"  
/sex="male"  
/tissue\_type="retina"  
/dev\_stage="55 year old"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares retina N2b4HR"  
/notes="Organ: eye; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCCCGCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian and total cellular poly(A)<sup>+</sup> RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Donaldso."

## ORIGIN

## Alignment Scores:

Pred. No.: 5.61e-68 Length: 681  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BX118596 (1-681)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
Db 243 TCCCTCTTCGCGACCGTCTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGACAA 302  
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
Db 303 TTGCTCATCTCTCTCGATGTGAAGCACTTCTCCCGAGGACCTCACCGTGAAGGTGCAG 362  
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60  
Db 363 GACGACTTTGTGGAGATCCACGGAAGCAACAGCGCGCAGACGACGCTACATT 422  
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
Db 423 TCCCGTGAGTTCCACCGCGCTTACCGCTCGCTCCACGTGGACCGACCTCTCTCT 482  
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
Db 483 TGCTCCCTGTCTGGCGATGGCATGTGACCTTCTGTGGCCCAAGATCCAGACTGGCGCTG 542  
QY 101 AspAlaThrHisAlaGluAlaIleProValSerArgGluGluLysProThrSerAla 120  
Db 543 GATGCCACCCACCGCGAGCGACCATCCCGTGTCTGCGGGAGGAGACCCCACTCGGCT 602  
QY 121 ProSerSer 123  
Db 603 CCTCTGCTCC 611

## RESULT 35

## BF727324

## LOCUS

DEFINITION BF727324 698 bp mRNA linear EST 05-JAN-2001  
sapiens cDNA clone by19h12 5', mRNA sequence.

## ACCESSION

## BF727324

## VERSION

## BF727324.1

## KEYWORDS

## EST.

## SOURCE

## Homo sapiens (human)

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.  
NEIBANK: EST analysis and bioinformatics for ocular genomics  
Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: grame@helix.nih.gov  
Plate: 19 row: h column: 12  
Seq primer: M13RP1 reverse primer (ABI).  
Location/Qualifiers  
1. .698  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="by19h12"  
/tissue\_type="Lens"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human Lens cDNA (Un-normalized, unamplified):  
By"

## FEATURES

## source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="by19h12"  
/tissue\_type="Lens"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human Lens cDNA (Un-normalized, unamplified):  
By"

/notes="Organ: Eye; Vector: pCMVSPORT6; Two human lenses  
from different adults (both approximately 40 years old)  
together yielded 20ug of total RNA and 150ng mRNA for cDNA  
library synthesis. A directionally cloned cDNA library in  
the pCMVSPORT6 vector was constructed at Life  
Technologies, essentially following the protocols of the  
SuperScript Plasmid System full details of which are  
contained in the manufacturer's instruction manual  
(http://www.lifetech.com/). First strand synthesis was  
carried out using a Not I primer-adaptor  
[5'-pOACTAGTCTTAGATCGGCGCGCC(T)15-3']. Not I/blunt  
end inserts were cloned into the Not I/EcoR V sites in the  
vector. EST analysis was performed on the unamplified  
library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

## Alignment Scores:

Pred. No.: 5.8e-68 Length: 698  
Score: 639.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF727324 (1-698)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
Db 205 TCCCTCTTCGCGACCGTCTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGACAA 264  
QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40  
Db 265 TTGCTCATCTCTCTCGATGTGAAGCACTTCTCCCGAGGACCTCACCGTGAAGGTGCAG 324  
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60  
Db 325 GACGACTTTGTGGAGATCCACGGAAGCAACAGCGCGCAGACGACCTACATT 384  
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
Db 385 TCCCGTGAGTTCCACCGCGCTTACCGCTCGCTCCACGTGGACCGACCTCTCTCT 444  
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
Db 445 TGCTCCCTGTCTGCGATGGCATGTGACTTCTGTGGCCCAAGATCCAGATCGGCTG 504



/dev\_stage="fetal"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-E00"  
 /note="Organ: eye; Vector: pTT3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-E00 is a cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGGATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,35e-67 Length: 559  
 Score: 635.00 Matches: 122  
 Percent Similarity: 99.19% Conservative: 0  
 Best Local Similarity: 99.19% Mismatches: 1  
 Query Match: 99.37% Indels: 0  
 DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM722779 (1-559)

QY 1 SerLeuPheArGThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 DB 186 TCCCTCTTCCGACCGTCTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGCAAG 245  
 QY 21 PheValIlePheLeuAspValIlyHisPheSerProGluAspLeuThrValIlyValGln 40  
 DB 246 TTGTCATCTTCTCGATGTAAGCACTTCTCCCGGAGGACCTCACCGTGAAGTGCAG 305  
 QY 41 AspAspPheValGluIleHisGlyIlyHisAsnGluArgGlnAspAspHisGlyTyrlle 60  
 DB 306 GACGACTTTGTGGAGATCCACGGAAGCAACAGCGCCGAGGACGACGACGCTACATT 365  
 QY 61 SerArgGluPheHisArgArgTyrrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 DB 366 TCCCGTGAAGTTCACCGCGCTACCGCTCGCGTCCCACTGCGGACGCTCGCCCTCTCT 425  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProIlyValIleGlnThrGlyLeu 100  
 DB 426 TGCTCCCTGTCGCCGATGCGATGCTGACCTTCTGTGGCCCAAGATCCAGCTGGCTG 485  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluIlyProThrSerAla 120  
 DB 486 GATGCCACCCAGCGGAGCGCCATCCCGTGTCCGCGGAGGAGAGCCACCTCGGNT 545  
 QY 121 ProSerSer 123  
 DB 546 CCCTCGTCC 554

## RESULT 38

## BF726358

## LOCUS

## DEFINITION

## BY05d12.y1 Human Lens cDNA (Un-normalized, amplified); BY Homo

## sapiens cDNA clone by05d12 5', mRNA sequence.

## BF726358

## VERSION

## BF726358.1

## KEYWORDS

## EST

## SOURCE

## ORGANISM

## Homo sapiens

## Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## 1 (bases 1 to 532)

## Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.

## NEIBANK: EST analysis and bioinformatics for ocular genomics

## JOURNAL

## COMMENT

Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
 Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: grame@helix.nih.gov  
 Plate: 95 row: d column: 12  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1. .532

## FEATURES

## Source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="by05d12"  
 /tissue\_type="Lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens cDNA (Un-normalized, amplified)"  
 By"

/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACGATGTTAGATCGGAGCGCGCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,22e-67 Length: 532  
 Score: 633.00 Matches: 122  
 Percent Similarity: 99.19% Conservative: 0  
 Best Local Similarity: 99.19% Mismatches: 1  
 Query Match: 99.06% Indels: 0  
 DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF726358 (1-532)

QY 1 SerLeuPheArGThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20  
 DB 152 TCCCTCTTCCGACCGTCTGGACTCCGGCATCTCTGAGGTTCGATCCGACCGGCAAG 211  
 QY 21 PheValIlePheLeuAspValIlyHisPheSerProGluAspLeuThrValIlyValGln 40  
 DB 212 TTGTCATCTTCTCGATGTAAGCACTTCTCCCGGAGGACCTCACCGTGAAGTGCAG 271  
 QY 41 AspAspPheValGluIleHisGlyIlyHisAsnGluArgGlnAspAspHisGlyTyrlle 60  
 DB 272 GACGACTTTGTGGAGATCCACGGAAGCAACAGCGCCGAGGACGACGACGCTACATT 331  
 QY 61 SerArgGluPheHisArgArgTyrrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 DB 332 TCCCGTGAAGTTCACCGCGCTACCGCTCGCGTCCCACTGCGGACGACGCTACATT 391  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProIlyValIleGlnThrGlyLeu 100  
 DB 392 TGCTCCCTGTCGCCGATGCGATGCTGACCTTCTGTGGCCCAAGATCCAGCTGGCTG 451  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluIlyProThrSerAla 120  
 DB 452 GATGCCACCCAGCGGAGCGCCATCCCGTGTCCGCGGAGGAGAGCCACCTCGGCT 511  
 QY 121 ProSerSer 123  
 DB 511

Db 512 CCCTCGTCC 520

RESULT 39  
AY419530  
LOCUS  
DEFINITION  
Pan troglodytes CRYAA gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
AY419530  
ACCESSION  
AY419530  
VERSION  
AY419530.1 GI:39775487  
KEYWORDS  
GSS.  
SOURCE  
Pan troglodytes (chimpanzees)  
ORGANISM  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
REFERENCE  
1 (bases 1 to 522)  
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL  
Science 302 (5652), 1960-1963 (2003)  
PUBMED  
14671302  
REFERENCE  
2 (bases 1 to 522)  
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE  
Direct Submision  
JOURNAL  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT  
These sequences were made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
Location/Qualifiers  
source  
1..522  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>522  
/gene="CRYAA"  
/locus\_tag="HCW6921"

gene

ORIGIN

Alignment Scores:  
Pred. No.: 2.88e-67 Length: 522  
Score: 632.00 Matches: 122  
Percent Similarity: 99.19% Conservative: 0  
Best Local Similarity: 99.19% Mismatches: 1  
Query Match: 98.90% Indels: 0  
DB: 9 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x AY419530 (1-522)

Qy 1 SerLeuPheArgThrValLeuAspSerGlyLeuValArgSerAspArgAspLys 20  
Db 151 TCCCTCTTTCCGACCGCTGCGACTCCGCGCATCTCTGAGTTTCGATCCGCGGACAAG 210  
Qy 21 PheValIlePheLeuAspValIleHisGlyValHisAsnGluArgGlnAspHisGlyTyrIle 40  
Db 211 TTGTCATCTTCTCGATGTGAAGACATCTCCCGGAGGACCTCACCCTGAAGGTGCAG 270  
Qy 41 AspAspPheValGluIleHisGlyValHisAsnGluArgGlnAspHisGlyTyrIle 60  
Db 271 GACGACTTTGTGGAGATCCACGGAAGACACACGAGCGCCAGNNNGACCGCTACATT 330  
Qy 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
Db 331 TCCGTCGATGTTCCACCGCGCTTACCGCTCCACGTCGAGCAGTCCGCGCTCTCT 390  
Qy 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100  
Db 391 TGCTCCCTGTCCGCGGATGGCATGCTGACCTTCTGTGTGCCCCCAAGATCCAGATGGCCTG 450

Qy 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120  
Db 451 GATGCCACCCAGCCGAGCGCCATCCCGTGTGCGGAGGAGAGAGCCACCTCGGCT 510  
Qy 121 ProSerSer 123  
Db 511 CCCTCGTCC 519  
RESULT 40  
BM696477  
LOCUS  
DEFINITION  
UI-E-DW0-agj-1-20-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone  
UI-E-DW0-agj-1-20-0-UI 5', mRNA sequence.  
BM696477  
ACCESSION  
BM696477  
VERSION  
BM696477.1 GI:19009735  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 519)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL  
Genome Res. 6 (9), 791-806 (1996)  
MEDLINE  
97044477  
PUBMED  
8889548  
COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1..519  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="UI-E-DW0-agj-1-20-0-UI"  
/tissue\_type="lens"  
/dev\_stages="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-DW0"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
UI-E-DW0 is a cDNA library containing the following  
tissues(s): lens. The library was constructed according to  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was ligated to an EcoR I adaptor, digested with Not  
I, and cloned directionally into pT7T3-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is CGATTAGCGA. This library  
was created for the program, Gene Discovery in the Visual  
System, supported by National Eye Institute (NEI)."

ORIGIN

Alignment Scores:  
Pred. No.: 3.79e-67 Length: 519  
Score: 631.00 Matches: 122  
Percent Similarity: 99.19% Conservative: 0  
Best Local Similarity: 99.19% Mismatches: 1

Query Match: 98.75% Indels: 0  
 DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM696477 (1-519)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspIlys 20  
 DB 135 TCCCTCTTCGCGACCGTCTGGACTCCGGCATCTCTGAGTTCCATCCGACCGGCAAG 194  
 QY 21 PheValIlePheLeuAspValIysHisPheSerProGluAspLeuThrValIysValGln 40  
 DB 195 TTCGTCATCTTCTCCGTGTAAGCACTTCTCCCGGAGGACCTCACCGTGAAGTGCAG 254  
 QY 41 AspAspPheValGluIleHisGlyIysHisAsnGluArgGlnAspAspHisGlyIle 60  
 DB 255 GACGACTTTGTGGAGATCCACGGAAGCAACAGCAGCGCCAGGACGACCGGTACATT 314  
 QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80  
 DB 315 TCCCGTGAGTTCCACCGCGCTACCGCTGCGTCCAACTGGACCACTCGGCCCTCTCT 374  
 QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProIysIleGlnThrGlyLeu 100  
 DB 375 TGCTCCCTGTCTGCGATGGATGCTGACCTTCTGTGGCCCAAGATCCAGACTGGCCTG 434  
 QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluIysProThrSerAla 120  
 DB 435 GATGCCACCCACGCGGAGGACCATCCCGTGTGCGGGGAGGAGAACCCCTCGGCT 494  
 QY 121 ProSerSer 123  
 DB 495 CCCTCGTCC 503

## RESULT 41

BF726253

LOCUS

DEFINITION

by03h05.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo

ACCSSION

BF726253

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

## FEATURES

source

1..577  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="by03h05"  
 /tissue\_type="Lens"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Lens cDNA (Un-normalized, unamplified):  
 By"  
 /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses  
 from different adults (both approximately 40 years old)  
 together yielded 20ug of total RNA and 150ng mRNA for cDNA  
 library synthesis. A directionally cloned cDNA library in

the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (<http://www.lifetech.com/>). First strand synthesis was carried out using a Not I primer-adaptor [5'-TGACTAGTTCTAGATCGGAGCGCCGCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

## Alignment Scores:

Pred. No.: 5,78e-67 Length: 577  
 Score: 630.00 Matches: 122  
 Percent Similarity: 99.19% Conservative: 0  
 Best Local Similarity: 99.19% Mismatches: 1  
 Query Match: 98.59% Indels: 0  
 DB: 2 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BF726253 (1-577)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspIlys 20

DB 198 TCCCTCTTCGCGACCGTCTGGACTCCGGCATCTCTGAGTTCCATCCGACCGGCAAG 257

QY 21 PheValIlePheLeuAspValIysHisPheSerProGluAspLeuThrValIysValGln 40

DB 258 TTCGTCATCTTCTCCGTGTAAGCACTTCTCCCGGAGGACCTCACCGTGAAGTGCAG 317

QY 41 AspAspPheValGluIleHisGlyIysHisAsnGluArgGlnAspAspHisGlyIle 60

DB 318 GACGACTTTGTGGAGATCCACGGAAGCAACAGCAGCGCCAGGACGACCGGTACATT 377

QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80

DB 378 TCCCGTGAGTTCCACCGCGCTACCGCTGCGTCCAACTGGACCACTCGGCCCTCTCT 437

QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProIysIleGlnThrGlyLeu 100

DB 438 TGCTCCCTGTCTGCGATGGATGCTGACCTTCTGTGGCCCAAGATCCAGACTGGCCTG 497

QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluIysProThrSerAla 120

DB 498 GATGCCACCCACGCGGAGGACCATCCCGTGTGCGGGGAGGAGAACCCCTCGGCT 557

QY 121 ProSerSer 123

DB 558 CCCTCGTCC 566

## RESULT 42

BF726422

LOCUS

DEFINITION

by06d05.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo

ACCSSION

BF726422

VERSION

BF726422.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

587 bp mRNA linear

EST 05-JAN-2001

sapiens cDNA clone by06d05 5', mRNA sequence.

EST.

GI:12042333

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 587)

Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.

NEIBANK: EST analysis and bioinformatics for ocular genomics

Invest. Ophthalmol. Vis. Sci. 41 (2000) In press

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

```
Plate: 06 row: d column: 05
Seq primer: M13RP1 reverse primer (ABI).
FEATURES
  source
    1..587
      Location/Qualifiers
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="by06d05"
        /tissue_type="Lens"
        /dev_stage="Adult"
        /lab_host="EMDH10B"
        /clone_lib="Human Lens cDNA (Un-normalized, unamplified);
        By"
        /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
        from different adults (both approximately 40 years old)
        together yielded 20ug of total RNA and 150ng mRNA for cDNA
        library synthesis. A directionally cloned cDNA library in
        the pCMVSPORT6 vector was constructed at Life
        Technologies, essentially following the protocols of the
        SuperScript Plasmid System full details of which are
        contained in the manufacturer's instruction manual
        (http://www.lifetech.com/). First strand synthesis was
        carried out using a Not I primer-adaptor
        [5'-pGACTAGTCTAGATCGGCGGCCGCC(T)15-3']. Not I/blunt
        end inserts were cloned into the Not I/EcoR V sites in the
        vector. EST analysis was performed on the unamplified
        library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN
Alignment Scores:
Pred. No.:      5,91e-67      Length:      587
Score:          630.00      Matches:     122
Percent Similarity: 99.19%      Conservative: 0
Best Local Similarity: 99.19%      Mismatches:  1
Query Match:    98.59%      Indels:      0
DB:             2          Gaps:         0

US-10-657-740-1_COPY_51_173 (1-123) x BF726422 (1-587)
Qy      1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db      158 TCCCTCTTCCGACCGCTGCTGACCTCCGCGCATCTCTGAGTTCGATCCGCGGACGAC 217
Qy      21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40'
Db      218 TTCGTTCATCTTCTCGATGTGAAGCACCTTCTCCCGGAGGACCTCACCGTGAAGGTGCG 277
Qy      41 AspAspPheValGluIleHisGlyIleHisGlyIleHisGlnArgGlnAspAspHisGlyTyrIle 60
Db      278 GACGACTTTGTGGAGATCCAGGAAGCACACACGAGCGCCAGGACGACCGGCTACATTT 337
Qy      61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db      338 TCCCGTGAGTTCACCGCGCTACCGCTTCGCTCCAGCTGAGGACGACGCTCGGCTCTCT 397
Qy      81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
Db      398 TGCTCCCTGTCTGCCGATGGCATGCTGACCTTCTGTGGCCCCCAAGATCCAGACTGGCCTG 457
Qy      101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db      458 GATGCCACCCACGCCGAGCGGCCATCCNCGTGTGCGCGGAGGAGAGGCCACCTCGGCT 517
Qy      121 ProSerSer 123
Db      518 CCTCGTCC 526

RESULT 43
LOCUS      BF726330
DEFINITION by05b01.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo
           sapiens cDNA clone by05b01 5', mRNA sequence.
ACCESSION  BF726330
```

```
VERSION      BF726330.1      GI:12042241
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 597)
AUTHORS      Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
TITLE        NEIBANK: EST analysis and bioinformatics for ocular genomics
JOURNAL      Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT      Contact: Wistow G
              Section on Molecular Structure and Function
              National Eye Institute
              6/331, NIH, Bethesda, MD 20892-2740, USA
              Tel: 301 402 3452
              Fax: 301 496 0078
              Email: graeme@helix.nih.gov
              Plate: 05 row: b column: 01
              Seq primer: M13RP1 reverse primer (ABI).
FEATURES     source
              Location/Qualifiers
                1..597
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="by05b01"
                  /tissue_type="Lens"
                  /dev_stage="Adult"
                  /lab_host="EMDH10B"
                  /clone_lib="Human Lens cDNA (Un-normalized, unamplified);
                  By"
                  /note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
                  from different adults (both approximately 40 years old)
                  together yielded 20ug of total RNA and 150ng mRNA for cDNA
                  library synthesis. A directionally cloned cDNA library in
                  the pCMVSPORT6 vector was constructed at Life
                  Technologies, essentially following the protocols of the
                  SuperScript Plasmid System full details of which are
                  contained in the manufacturer's instruction manual
                  (http://www.lifetech.com/). First strand synthesis was
                  carried out using a Not I primer-adaptor
                  [5'-pGACTAGTCTAGATCGGCGGCCGCC(T)15-3']. Not I/blunt
                  end inserts were cloned into the Not I/EcoR V sites in the
                  vector. EST analysis was performed on the unamplified
                  library at the NIH Intramural Sequencing Center (NISC)."
ORIGIN
Alignment Scores:
Pred. No.:      8,02e-67      Length:      597
Score:          629.00      Matches:     121
Percent Similarity: 98.37%      Conservative: 0
Best Local Similarity: 98.37%      Mismatches:  2
Query Match:    98.44%      Indels:      0
DB:             2          Gaps:         0

US-10-657-740-1_COPY_51_173 (1-123) x BF726330 (1-597)
Qy      1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
Db      152 TCCCTCTTCCGACCGCTGCTGACCTCCGCGCATCTCTGAGTTCGATCCGCGGACGAC 211
Qy      21 PheValIlePheLeuAspValIleHisPheSerProGluAspLeuThrValIleValGln 40
Db      212 TTCGTTCATCTTCTCGATGTGAAGCACCTTCTCCCGGAGGACCTCACCGTGAAGGTGCG 271
Qy      41 AspAspPheValGluIleHisGlyIleHisGlyIleHisGlnArgGlnAspAspHisGlyTyrIle 60
Db      272 GACGACTTTGTGGAGATCCAGGAAGCACACACGAGCGCCAGGACGACCGGCTACATTT 331
Qy      61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
Db      332 TCCCGTGAGTTCACCGCGCTACCGCTTCGCTCCAGCTGAGGACGACGCTCGGCTCTCT 391
Qy      81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu 100
```

```

Db      392  TGTCTCCTGTCTGCGGATGCTGATCTTCTGTGCCCCAAGATCCAGACTGGCGCTG 451
Qy      101  AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAla 120
Db      452  GATGCCACCCAGCCGAGGAGCCATCCCGTGTGCGGGAGGAGAGCCACCTCGGCT 511
Qy      121  ProSerSer 123
Db      512  CCTCTGTC 520

```

```

RESULT 44
LOCUS   BQ640267
DEFINITION
Homo sapiens cDNA (Un-normalized, unamplified): hd/he
ACCESSION BQ640267
VERSION   BQ640267.1 GI:21764726
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 604)
AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A.,
          Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE     Expressed sequence tag analysis of human retina for the NEIBank
          Project: Retbindin, an abundant, novel retinal cDNA and alternative
          splicing of other retina-preferred gene transcripts
JOURNAL   Mol. Vis. 8 (4), 196-204 (2002)
MEDLINE   22103461
PUBMED    12107411
COMMENT   Contact: Wistow G
          Section on Molecular Structure and Function
          National Eye Institute
          6/331, NIH, Bethesda, MD 20892-2740, USA
          Tel: 301 402 3452
          Fax: 301 496 0078
          Email: graeme@helix.nih.gov
          Plate: 26 row: c column: 10
          Seq primer: M13RPL reverse primer (ABT).
          Location/Qualifiers
            1..604
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="he26c10"
              /tissue_type="Retina"
              /dev_stage="Adult"
              /lab_host="BMDH10B"
              /clone_lib="Human Retina cDNA (Un-normalized,
              unamplified): hd/he"
              /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
              was dissected from two 80 year old donors with no observed
              eye disease. 100ug of total RNA was used for library
              construction. A directionally cloned cDNA library in the
              pSPORT1 vector (Life Technologies) was constructed at
              Bioserve Biotechnology (Laurel MD) essentially following
              the protocols of the SuperScript Plasmid System full
              details of which are contained in the manufacturer's
              instruction manual (http://www.lifetech.com/). First
              strand synthesis was carried out using a Not I
              primer-adaptor
              [5'-pGACTAGTCTTAGATCGGAGCGCGCC(T)15-3']. EST analysis
              was performed on the unamplified library at the NIH
              Intramural Sequencing Center (NISC)."
```

```

FEATURES
source
1..604
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

```

FEATURES
source
1..580
BM706139
UI-B-DW0-agg-b-16-0-UI-r1 UI-B-DW0 Homo sapiens cDNA clone
BM706139
UI-B-DW0-agg-b-16-0-UI 5', mRNA sequence.
BM706139.1 GI:19019397
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@iowa.edu
          Tissue Procurement: Dr. Gregg Hageman
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing Arrayed by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Dr. M. Bento Soares, University of Iowa
          Genetics (www.resgen.com).
          Seq primer: M13 Reverse.
          Location/Qualifiers
            1..580
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="UI-B-DW0-agg-b-16-0-UI"
              /tissue_type="lens"
```

## ORIGIN

Alignment Scores:

```

Pred. No.:      8.15e-67      Length:      604
Score:          629.00      Matches:      121
Percent Similarity: 98.37%      Conservative: 0
Best Local Similarity: 98.37%      Mismatches: 2

```



```

/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DW0"
/Note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-DW0 is a cDNA library containing the following
tissue(s): lens. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGATTAGGCA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 1.81e-66 Length: 580
Score: 626.00 Matches: 120
Percent Similarity: 99.17% Conservative: 0
Best Local Similarity: 99.17% Mismatches: 1
Query Match: 97.97% Indels: 0
DB: 4 Gaps: 0
```

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM706139 (1-580)

```

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
DB 206 TCCCTCTTCCGACCGTGGACCTCCGACATCTCTGAGGTTCGATCCGACCGGACAAG 265
QY 21 PheValIlePheLeuAspVallyHisPheSerProGluAspLeuThrVallyValGln 40
DB 266 TTGCTCATCTCTCGATGTGAGACATCTCTCCCGAGGAGCTCACCCTGAGGTGAG 325
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60
DB 326 GACGACTTTGTGGAGATCCAGGAAAGCACACAGAGCGCCAGGACGACCGGCTACATT 385
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
DB 386 TCCCGTGAGTTCCACCGCGCTTACCGCTCCGCTCAACAGTGGACGACGCTGCTCT 445
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysValIleGlnThrGlyLeu 100
DB 446 TGCTCNCCTGTGCGGATGGCATGTGACCTTCTGTGGCCCCAAGATCCAGACTGGGCTG 505
QY 101 AspAlaThrHisAlaGluArgAlaIleProValSerArgGluLysProThrSerAla 120
DB 506 GATGCCACCCAGCGGAGGACCATCCCGTGTCCGGGAGGAGAGCCCACTCGGCT 565
QY 121 Pro 121
DB 566 CCC 568
```

## RESULT 46

```

BM722650
LOCUS BM722650 586 bp mRNA linear EST 01-MAR-2002
DEFINITION UI-E-E00-ahy-a-10-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone
UI-E-E00-ahy-a-10-0-UI 5', mRNA sequence.
ACCESSION BM722650
VERSION BM722650.1 GI:19043377
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 586)
AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
```

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477

Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).

Seq primer: M13 Reverse.

## FEATURES

## source

```

1..586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-E00-ahy-a-10-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-E00"
/Note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-E00 is a cDNA library containing the following
tissue(s): fetal eye. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGCGTATACC. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
```

## ORIGIN

```

Alignment Scores:
Pred. No.: 1.83e-66 Length: 586
Score: 626.00 Matches: 123
Percent Similarity: 99.19% Conservative: 0
Best Local Similarity: 99.19% Mismatches: 0
Query Match: 97.97% Indels: 1
DB: 4 Gaps: 0
```

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM722650 (1-586)

```

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20
DB 200 TCCCTCTTCCGACCGTGGACCTCCGACATCTCTGAGGTTCGATCCGACCGGACAAG 259
QY 21 PheValIlePheLeuAspVallyHisPheSerProGluAspLeuThrVallyValGln 40
DB 260 TTGCTCATCTCTCGATGTGAGACATCTCTCCCGAGGAGCCTCACCCTGAGGTGAG 319
QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle 60
DB 320 GACGACTTTGTGGAGATCCAGGAAAGCACACAGGCGCCAGGACGACCGGCTACATT 379
QY 61 SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer 80
DB 380 TCCCGTGAGTTCCACCGCGCTTACCGCTCCGCTCAACAGTGGACGACGCTGCTCTCT 439
QY 81 CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysValIleGlnThrGlyLeu 100
```

Db 440 TGCTCCCTGTCCTCCGATGCGATGACCTTCTGTGGCCCCAAGATCCAGACTGGCTG 499

QY 101 AspAlaThr-HisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAl 120

Db 500 GATGCCACCCACGCCGAGGAGCCATCCCTGTCGCGGAGGAGAGCCACCTCGGC 559

QY 120 aProSerSer 123

Db 560 TCCTCGTCC 569

RESULT 47

LOCUS BM696670

DEFINITION UI-E-DW0-agk-o-16-0-UI.r1 mRNA linear EST 28-FEB-2002

ACCESSION UI-E-DW0-agk-o-16-0-UI 5', mRNA sequence.

VERSION BM696670

KEYWORDS EST.

SOURCE BM696670.1 GI:19009928

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 607)

TITLE Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalizaton and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Reverse.

## FEATURES

source

1. .607

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-E-DW0-agk-o-16-0-UI"

/tissue\_type="lens"

/dev\_stage="adult"

/lab\_host="PH108 (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-E-DW0"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGATTACGGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

## ORIGIN

Alignment Scores:

Pred. No.: 1.92e-66 Length: 607

Score: 626.00 Matches: 123

Percent Similarity: 99.19% Conservative: 0

Best Local Similarity: 99.19% Mismatches: 0

Query Match: 97.97% Indels: 1

DB: 4 Gaps: 0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM696670 (1-607)

QY 1 SerLeuPheArgThrValLeuAspSerGlyIleSerGluValArgSerAspArgAspLys 20

Db 84 TCCTCTTCGACCCGCTGCTGACTCCGGCATCTCTGAGTTCCGATCCGCGACAAG 143

QY 21 PheValIlePheLeuAspValLysHisPheSerProGluAspLeuThrValLysValGln 40

Db 144 TTTCGTCACTTCTCTCGATGTGAAGCACATCTCCCGGAGGACCTCACCTGAAGTGCAG 203

QY 41 AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTrile 60

Db 204 GACGACTTTGTGGAGATCCACGAAAGCACAAACAGCGCCAGGACGACCGCTACATT 263

QY 61 SerArgGlu-PheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSe 80

Db 264 TCCCGTGAAGTTCCACCGCCGCTACCGCTGCCCTCCAACTGACGAGGAGGAGGAGGAG 323

QY 80 rCysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLe 100

Db 324 TTGCTCCCTGTCCTCCGATGCGATGCTGACCTTCTGTGGCCCCAAGATCCAGACTGGCCT 383

QY 100 uAspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSerAl 120

Db 384 GGATGCCACCCACCGCGAGCGAGCCATCCCTGTCGCGGAGGAGGAGGAGGAGGAGGAG 443

QY 120 aProSerSer 123

Db 444 TCCTCGTCC 453

## RESULT 48

LOCUS BM706160

DEFINITION UI-E-DW0-agg-f-12-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone

ACCESSION UI-E-DW0-agg-f-12-0-UI 5', mRNA sequence.

VERSION BM706160

KEYWORDS EST.

SOURCE BM706160.1 GI:19019418

ORGANISM Homo sapiens (human)

## REFERENCE

AUTHORS 1 (bases 1 to 512)

TITLE Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

## COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Reverse.

## FEATURES

source

Location/Qualifiers

1. .512

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"



UI-E-DW0-agl-k-14-0-UI 5', mRNA sequence.

ACCESSION  
BM696853  
VERSION  
BM696853.1 GI:19010111  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS  
TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 498)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene  
discovery

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Genome Res. 6 (9), 791-806 (1996)  
97044477

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MBBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1..498

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-E-DW0-agl-k-14-0-UI"

/tissue\_type="lens"

/dev\_stage="adult"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-E-DW0"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-DW0 is a cDNA library containing the following

tissue(s): lens. The library was constructed according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. First strand cDNA synthesis was primed with an

oligo-dT primer containing a Not I site. Double stranded

cDNA was ligated to an EcoR I adaptor, digested with Not

I, and cloned directionally into pT7T3-Pac vector. The

oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tag for this library is CGATTAGCGA. This library

was created for the program, Gene Discovery in the Visual

System, supported by National Eye Institute (NEI)."

# ORIGIN

## Alignment Scores:

Pred. No.:	2,51e-65	Length:	498
Score:	616.00	Matches:	118
Percent Similarity:	99.16%	Conservative:	0
Best Local Similarity:	99.16%	Mismatches:	1
Query Match:	96.40%	Indels:	0
DB:	4	Gaps:	0

US-10-657-740-1\_COPY\_51\_173 (1-123) x BM696853 (1-498)

QY 1 SerLeuPheArgThrValLeuAspSerGlyLeSerGluValArgSerAspArgLys 20

Db 140 TCCCTCTTCCGACCGCTGCTCGGACATCTCGAGTTTCGATCCGACCGGACAG 199

QY 21 PheValIlePheLeuAspValIysHisPheSerProGluaspLeuThrValIysValGln 40

Db 200 TTGTCATCTTCTTCGATGTGAGACATCTTCCCGGAGGACCTCACCGTGAAGGTGCAG 259

QY	41	AspAspPheValGluIleHisGlyLysHisAsnGluArgGlnAspAspHisGlyTyrIle	60
Db	260	GACGACTTTGTGGAGATCCACGAAAGCACAAACGAGCGCCAGGACGACACCGCTACATT	319
QY	61	SerArgGluPheHisArgArgTyrArgLeuProSerAsnValAspGlnSerAlaLeuSer	80
Db	320	TCCCGTGAGTTCCACCGCGCTACCGCTCCAAACGTGGACCATCGGNCCTCTCT	379
QY	81	CysSerLeuSerAlaAspGlyMetLeuThrPheCysGlyProLysIleGlnThrGlyLeu	100
Db	380	TGCTCCCTGCTCGCCGATGGATGCTGACCTTCTGTGCCCCCAGATCCAGACTGGCCTG	439
QY	101	AspAlaThrHisAlaGluArgAlaIleProValSerArgGluGluLysProThrSer	119
Db	440	GATGCCACCCACGCGAGCGACCATCCCCCTGTTCGGGGAGGAGAGCCACCTCG	496

Search completed: May 30, 2005, 08:48:01

Job time : 2324.47 secs